GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 14:33:19 ; Search time 25.75 Seconds
(without alignments)
36.985 Million cell updates/sec

Title: US-09-881-490-150
Perfect score: 31
Sequence: 1 LIQLFH 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	31	100.0	6	17	AAW04061	Antifungal peptide
2	31	100.0	6	18	AAW44549	Anti-fungal peptid
3	31	100.0	6	20	AAY00526	Antifungal peptide
4	31	100.0	6	22	AAB65450	Anti-fungal peptid
5	31	100.0	7	17	AAW04022	Antifungal peptide
6	31	100.0	7	17	AAW04023	Antifungal peptide
7	31	100.0	7	18	AAW44541	Anti-fungal peptid
8	31	100.0	7	18	AAW44542	Anti-fungal peptid
9	31	100.0	7	18	AAW43731	Bactericidal/perme
10	31	100.0	7	18	AAW43732	Bactericidal/perme
11	31	100.0	7	20	AAY00519	Antifungal peptide
12	31	100.0	7	20	AAY00518	Antifungal peptide
13	31	100.0	7	22	AAB65442	Anti-fungal peptid
14	31	100.0	7	22	AAB65443	Anti-fungal peptid
15	31	100.0	8	17	AAW04017	Antifungal peptide
16	31	100.0	8	17	AAW04018	Antifungal peptide
17	31	100.0	8	17	AAW04019	Antifungal peptide
18	31	100.0	8	18	AAW44534	Anti-fungal peptid
19	31	100.0	8	18	AAW44535	Anti-fungal peptid
20	31	100.0	8	18	AAW44536	Anti-fungal peptid
21	31	100.0	8	18	AAW43724	Bactericidal/perme
22	31	100.0	8	18	AAW43725	Bactericidal/perme
23	31	100.0	8	18	AAW43726	Bactericidal/perme
24	31	100.0	8	20	AAY00511	Antifungal peptide
25	31	100.0	8	20	AAY00512	Antifungal peptide
26	31	100.0	8	20	AAY00513	Antifungal peptide
27	31	100.0	8	22	AAB65435	Anti-fungal peptid
28	31	100.0	8	22	AAB65436	Anti-fungal peptid
29	31	100.0	8	22	AAB65437	Anti-fungal peptid
30	31	100.0	9	17	AAW04011	Antifungal peptide
31	31	100.0	9	17	AAW04012	Antifungal peptide
32	31	100.0	9	17	AAW04013	Antifungal peptide
33	31	100.0	9	18	AAW43720	Bactericidal/perme
34	31	100.0	9	18	AAW43718	Bactericidal/perme
35	31	100.0	9	18	AAW43719	Bactericidal/perme
36	31	100.0	9	18	AAW44528	Anti-fungal peptid
37	31	100.0	9	18	AAW44529	Anti-fungal peptid
38	31	100.0	9	18	AAW44530	Anti-fungal peptid
39	31	100.0	9	20	AAY00505	Antifungal peptide
40	31	100.0	9	20	AAY00506	Antifungal peptide
41	31	100.0	9	20	AAY00507	Antifungal peptide
42	31	100.0	9	22	AAB68710	Peptide-based cons
43	31	100.0	9	22	AAB65429	Anti-fungal peptid
44	31	100.0	9	22	AAB65430	Anti-fungal peptid
45	31	100.0	9	22	AAB65431	Anti-fungal peptid

ALIGNMENTS

RESULT 1
AAW04061
ID AAW04061 standard; peptide; 6 AA.
XX
AC AAW04061;
XX
DT 04-NOV-1996 (first entry)
XX
DE Antifungal peptide XMP.317.
XX
KW Antifungal peptide; inhibitor; Domain III; polymorphonuclear leukocyte;
KW bactericidal/permeability-increasing protein; BPI; mammalian; PMN; fungi;
KW neutrophil; replication inhibitor; fungal infection; Aspergillus;
KW Cryptococcus; Candida; C.albicans; C.galabrat; C.krusei; C.lusitaniae;
KW C.parapsilosis; C.tropicalis; therapy.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 6
FT /note= "amidated"
XX
PN WO9608509-A1.
XX
PD 21-MAR-1996.
XX
PF 20-JUL-1995; 95WO-US09262.
XX
PR 13-JAN-1995; 95US-0372105.
PR 15-SEP-1994; 94US-0306473.
XX
PA (XOMA) XOMA CORP.

XX
PI Fadem MB, Lim E, Little RG;
XX
DR WPI; 1996-179900/18.
XX
PT Antifungal peptide(s) derived from Domain III of BPI protein - used
PT in vitro for killing or inhibiting replication of fungi, esp.
PT Candida species
XX
PS Claim 14; Page 150; 199pp; English.
XX
CC AAW04000-W04150 represent antifungal peptides. These sequences are
CC based on (or derived from) Domain III of the
CC bactericidal/permeability-increasing protein (BPI). BPI is a protein
CC that can be isolated from the granules of mammalian polymorphonuclear
CC leukocytes (PMNs or neutrophils). These antifungal peptides can be used
CC for killing, or inhibiting replication of, fungi in vitro. These
CC sequences can also be used for treatment of a fungal infection involving
CC fungi from the species Candida, Aspergillus and Cryptococcus. The
CC sequences are especially useful for treating C.albicans, C.galabrat,
CC C.krusei, C.lusitaniae, C.parapsilosis and C.tropicalis infections.
XX
SQ Sequence 6 AA;

Query Match 100.0%; Score 31; DB 17; Length 6;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LIQLFH 6
||||||
Db 1 LIQLFH 6

RESULT 2
AAW44549
ID AAW44549 standard; peptide; 6 AA.
XX
AC AAW44549;
XX
DT 27-APR-1998 (first entry)
XX
DE Anti-fungal peptide #150 based on BPI protein (residues 142-169).
XX
KW Anti-fungal peptide; bactericidal-permeability-increasing protein; BPI;
KW polymorphonuclear leukocyte; fungicide.
XX
OS Synthetic.
OS Mammalia.
XX
FH Key Location/Qualifiers
FT Modified-site 6
FT /note= "C-terminal amide"
XX
PN WO9704008-A1.
XX
PD 06-FEB-1997.
XX
PF 21-MAR-1996; 96WO-US03845.
XX
PR 20-JUL-1995; 95US-0504841.
XX
PA (XOMA) XOMA CORP.
XX
PI Fadem MB, Lim E, Little RG;
XX
DR WPI; 1997-132578/12.
XX
PT Anti-fungal peptide(s) derived from or based on domain III of
PT bactericidal-permeability-increasing protein - are used in vitro or
PT in vivo as a fungicides
XX
PS Claim 1; Page 187; 230pp; English.
XX
CC This is a specifically claimed anti-fungal peptide which is based on
CC domain III (amino acids 142-160) of bactericidal-permeability-increasing
CC protein (BPI), isolated from the granules of mammalian polymorphonuclear
CC leukocytes. It is used in compositions with diluents, carriers or
CC adjuvants to treat fungal infections in patients. It may also be used in
CC vitro to kill or inhibit the replication of fungi, such as in
CC decontaminating fluids and sterilising medical and implant devices.
XX
SQ Sequence 6 AA;

Query Match 100.0%; Score 31; DB 18; Length 6;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LIQLFH 6
||||||
Db 1 LIQLFH 6

RESULT 3
AAY00526
ID AAY00526 standard; Peptide; 6 AA.
XX
AC AAY00526;
XX
DT 07-MAY-1999 (first entry)
XX
DE Antifungal peptide XMP.317.
XX
KW Antifungal; BPI; bactericidal/permeability increasing protein;
KW Candida infection.
XX
OS Synthetic.
XX
PN US5858974-A.
XX
PD 12-JAN-1999.
XX
PF 21-MAR-1996; 96US-0621259.
XX
PR 21-MAR-1996; 96US-0621259.
PR 20-JUL-1995; 95US-0504841.
XX
PA (XOMA) XOMA CORP.
XX
PI Fadem MB, Lim E, Little RG;
XX
DR WPI; 1999-119956/10.
XX
PT Antifungal peptides - comprising part of bactericidal or
PT permeability-increasing protein sequence or related sequence
XX
PS Disclosure; Columns 157-158; 132pp; English.
XX
CC New peptides are provided which are based on Domain III (amino acids
CC 142-169) of human bactericidal/permeability-increasing protein (BPI).
CC The peptides all have a C-terminal amide. More particularly, the Claims
CC relate to: (1) a peptide that has an amino acid sequence of human BPI
CC from position 148 to position 161 (KSKVGWLIQLFHKK) and variants of the
CC sequence having antifungal activity; and (2) an antifungal peptide
CC having 6-14 amino acids comprising (a) a core sequence selected from
CC LIQL, IQLF, WLIQL, LIQLF and WLIQLF and (b) one or more cationic amino
CC acids selected from K, R, H, Orn (ornithine) and Dab (diaminobutyric
CC acid) at the N and/or C terminus of the core sequence. The new peptides
CC are used for killing or inhibiting replication of fungi in vitro; and
CC for treating fungal infections in vivo, in particular infections of
CC Candida, Aspergillus or Cryptococcus spp., especially C. albicans, C.
CC krusei, C. lusitaniae, C. parapsilosis or C. tropicalis. The peptide
CC can be administered topically, intravenously, orally or as an aerosol,
CC optionally together with a non-peptide antifungal agent.
XX
SQ Sequence 6 AA;

```
  Query Match             100.0%;  Score 31;  DB 20;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 9.3e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ||||||
Db          1 LIQLFH 6


RESULT 4
AAB65450
ID   AAB65450 standard; Peptide; 6 AA.
XX
AC   AAB65450;
XX
DT   27-MAR-2001  (first entry)
XX
DE   Anti-fungal peptide XMP.317.
XX
KW   Human; BPI; antifungal; polymorphonuclear leukocyte; neutrophil;
KW   bactericidal/permeability-increasing protein; bactericidal;
KW   fungal infection.
XX
OS   Homo sapiens.
XX
PN   US6156730-A.
XX
PD   05-DEC-2000.
XX
PF   08-JAN-1999;   99US-0227659.
XX
PR   21-MAR-1996;   96US-0621259.
PR   12-MAR-1993;   93US-0030644.
PR   15-JUL-1993;   93US-0093202.
PR   14-JAN-1994;   94US-0183222.
PR   11-MAR-1994;   94US-0209762.
PR   11-JUL-1994;   94US-0273540.
PR   15-SEP-1994;   94US-0306473.
PR   13-JAN-1995;   95US-0372105.
PR   20-JUL-1995;   95US-0504841.
XX
PA   (XOMA ) XOMA CORP.
XX
PI   Lim E,  Fadem MB,  Little RG;
XX
DR   WPI; 2001-090160/10.
XX
PT   Novel anti-fungal peptides derived from domain III of
PT   bactericidal/permeability-increasing protein useful for killing or
PT   inhibiting replication of fungi and for treating fungal infections  -
XX
PS   Example 2; Columns 163-164; 134pp; English.
XX
CC   The present invention relates to antifungal peptides (see
CC   AAB65301-B65550) derived from Domain III (amino acids 142-169) of
CC   bactericidal/permeability-increasing protein (BPI). The present sequence
CC   is one such antifungal peptide. BPI is a protein isolated from the
CC   granules of mammalian polymorphonuclear leukocytes (PMNs or
CC   neutrophils). BPI has potent bactericidal activity against a broad range
CC   of gram-negative bacteria. The peptides of the present invention are
CC   useful for killing or inhibiting replication of fungi, and treating
CC   infections caused by fungus belonging to Candida, Aspergillus,
CC   Cryptococcus species such as C.albicans, C.glabrata, C.krusei,
CC   C.lusitaniae, C.parapsilosis and C.tropicalis.
XX
SQ   Sequence   6 AA;

  Query Match             100.0%;  Score 31;  DB 22;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 9.3e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ||||||
Db          1 LIQLFH 6


RESULT 5
AAW04022
ID   AAW04022 standard; peptide; 7 AA.
XX
AC   AAW04022;
XX
DT   31-OCT-1996  (first entry)
XX
DE   Antifungal peptide XMP.309.
XX
KW   Antifungal peptide; inhibitor; Domain III; polymorphonuclear leukocyte;
KW   bactericidal/permeability-increasing protein; BPI; mammalian; PMN; fungi;
KW   neutrophil; replication inhibitor; fungal infection; Aspergillus;
KW   Cryptococcus; Candida; C.albicans; C.galabrat; C.krusei; C.lusitaniae;
KW   C.parapsilosis; C.tropicalis; therapy.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site    7
FT                    /note= "amidated"
XX
PN   WO9608509-A1.
XX
PD   21-MAR-1996.
XX
PF   20-JUL-1995;   95WO-US09262.
XX
PR   13-JAN-1995;   95US-0372105.
PR   15-SEP-1994;   94US-0306473.
XX
PA   (XOMA ) XOMA CORP.
XX
PI   Fadem MB,  Lim E,  Little RG;
XX
DR   WPI; 1996-179900/18.
XX
PT   Antifungal peptide(s) derived from Domain III of BPI protein  - used
PT   in vitro for killing or inhibiting replication of fungi, esp.
PT   Candida species
XX
PS   Claim 5; Page 147; 199pp; English.
XX
CC   AAW04000-W04160 represent antifungal peptides.  These sequences are
CC   based on (or derived from) Domain III of the
CC   bactericidal/permeability-increasing protein (BPI).  BPI is a protein
CC   that can be isolated from the granules of mammalian polymorphonuclear
CC   leukocytes (PMNs or neutrophils).  These antifungal peptides can be used
CC   for killing, or inhibiting replication of, fungi in vitro.  These
CC   sequences can also be used for treatment of a fungal infection involving
CC   fungi from the species Candida, Aspergillus and Cryptococcus.  The
CC   sequences are especially useful for treating C.albicans, C.galabrat,
CC   C.krusei, C.lusitaniae, C.parapsilosis and C.tropicalis infections.
XX
SQ   Sequence   7 AA;

  Query Match             100.0%;  Score 31;  DB 17;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 9.3e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ||||||
Db          2 LIQLFH 7


RESULT 6
AAW04023
ID   AAW04023 standard; peptide; 7 AA.
XX
AC   AAW04023;
```

```
XX
DT   31-OCT-1996  (first entry)
XX
DE   Antifungal peptide XMP.310.
XX
KW   Antifungal peptide; inhibitor; Domain III; polymorphonuclear leukocyte;
KW   bactericidal/permeability-increasing protein; BPI; mammalian; PMN; fungi;
KW   neutrophil; replication inhibitor; fungal infection; Aspergillus;
KW   Cryptococcus; Candida; C.albicans; C.galabrat; C.krusei; C.lusitaniae;
KW   C.parapsilosis; C.tropicalis; therapy.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   7
FT                   /note= "amidated"
XX
PN   WO9608509-A1.
XX
PD   21-MAR-1996.
XX
PF   20-JUL-1995;    95WO-US09262.
XX
PR   13-JAN-1995;    95US-0372105.
PR   15-SEP-1994;    94US-0306473.
XX
PA   (XOMA ) XOMA CORP.
XX
PI   Fadem MB,  Lim E,  Little RG;
XX
DR   WPI; 1996-179900/18.
XX
PT   Antifungal peptide(s) derived from Domain III of BPI protein  - used
PT   in vitro for killing or inhibiting replication of fungi, esp.
PT   Candida species
XX
PS   Claim 5; Page 147; 199pp; English.
XX
CC   AAW04000-W04160 represent antifungal peptides.  These sequences are
CC   based on (or derived from) Domain III of the
CC   bactericidal/permeability-increasing protein (BPI).  BPI is a protein
CC   that can be isolated from the granules of mammalian polymorphonuclear
CC   leukocytes (PMNs or neutrophils).  These antifungal peptides can be used
CC   for killing, or inhibiting replication of, fungi in vitro.  These
CC   sequences can also be used for treatment of a fungal infection involving
CC   fungi from the species Candida, Aspergillus and Cryptococcus.  The
CC   sequences are especially useful for treating C.albicans, C.galabrat,
CC   C.krusei, C.lusitaniae, C.parapsilosis and C.tropicalis infections.
XX
SQ   Sequence   7 AA;

  Query Match             100.0%;  Score 31;  DB 17;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 9.3e-05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ||||||
Db          1 LIQLFH 6


RESULT 7
AAW44541
ID   AAW44541 standard; peptide; 7 AA.
XX
AC   AAW44541;
XX
DT   27-APR-1998  (first entry)
XX
DE   Anti-fungal peptide #142 based on BPI protein (residues 142-169).
XX
KW   Anti-fungal peptide; bactericidal-permeability-increasing protein; BPI;
KW   polymorphonuclear leukocyte; fungicide.
XX
OS   Synthetic.
OS   Mammalia.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   7
FT                   /note= "C-terminal amide"
XX
PN   WO9704008-A1.
XX
PD   06-FEB-1997.
XX
PF   21-MAR-1996;    96WO-US03845.
XX
PR   20-JUL-1995;    95US-0504841.
XX
PA   (XOMA ) XOMA CORP.
XX
PI   Fadem MB,  Lim E,  Little RG;
XX
DR   WPI; 1997-132578/12.
XX
PT   Anti-fungal peptide(s) derived from or based on domain III of
PT   bactericidal-permeability-increasing protein - are used in vitro or
PT   in vivo as a fungicides
XX
PS   Claim 1; Page 184; 230pp; English.
XX
CC   This is a specifically claimed anti-fungal peptide which is based on
CC   domain III (amino acids 142-160) of bactericidal-permeability-increasing
CC   protein (BPI), isolated from the granules of mammalian polymorphonuclear
CC   leukocytes. It is used in compositions with diluents, carriers or
CC   adjuvants to treat fungal infections in patients. It may also be used in
CC   vitro to kill or inhibit the replication of fungi, such as in
CC   decontaminating fluids and sterilising medical and implant devices.
XX
SQ   Sequence   7 AA;

  Query Match             100.0%;  Score 31;  DB 18;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 9.3e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ||||||
Db          2 LIQLFH 7


RESULT 8
AAW44542
ID   AAW44542 standard; peptide; 7 AA.
XX
AC   AAW44542;
XX
DT   27-APR-1998  (first entry)
XX
DE   Anti-fungal peptide #143 based on BPI protein (residues 142-169).
XX
KW   Anti-fungal peptide; bactericidal-permeability-increasing protein; BPI;
KW   polymorphonuclear leukocyte; fungicide.
XX
OS   Synthetic.
OS   Mammalia.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   7
FT                   /note= "C-terminal amide"
XX
PN   WO9704008-A1.
XX
PD   06-FEB-1997.
XX
PF   21-MAR-1996;    96WO-US03845.
XX
PR   20-JUL-1995;    95US-0504841.
```

```
XX
PA   (XOMA ) XOMA CORP.
XX
PI   Fadem MB,  Lim E,  Little RG;
XX
DR   WPI; 1997-132578/12.
XX
PT   Anti-fungal peptide(s) derived from or based on domain III of
PT   bactericidal-permeability-increasing protein - are used in vitro or
PT   in vivo as a fungicides
XX
PS   Claim 1; Page 184; 230pp; English.
XX
CC   This is a specifically claimed anti-fungal peptide which is based on
CC   domain III (amino acids 142-160) of bactericidal-permeability-increasing
CC   protein (BPI), isolated from the granules of mammalian polymorphonuclear
CC   leukocytes. It is used in compositions with diluents, carriers or
CC   adjuvants to treat fungal infections in patients. It may also be used in
CC   vitro to kill or inhibit the replication of fungi, such as in
CC   decontaminating fluids and sterilising medical and implant devices.
XX
SQ   Sequence   7 AA;

  Query Match             100.0%;  Score 31;  DB 18;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 9.3e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ||||||
Db          1 LIQLFH 6


RESULT 9
AAW43731
ID   AAW43731 standard; peptide; 7 AA.
XX
AC   AAW43731;
XX
DT   20-APR-1998  (first entry)
XX
DE   Bactericidal/permeability increasing peptide XMP.309.
XX
KW   Bactericidal/permeability increasing peptide; BPI; fusion protein;
KW   bacterial infection; fungal infection; endotoxin; heparin;
KW   angiogenesis; fungicidal; recombinant DNA; vector.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   7
FT                   /note= "Amidated"
XX
PN   WO9735009-A1.
XX
PD   25-SEP-1997.
XX
PF   18-MAR-1997;   97WO-US05287.
XX
PR   22-MAR-1996;   96US-0621803.
XX
PA   (XOMA ) XOMA CORP.
XX
PI   Better MD;
XX
DR   WPI; 1997-480215/44.
XX
PT   Recombinant production of bactericidal/permeability increasing
PT   protein - by expression as a fusion protein in microbial host cells,
PT   then cleaving the BPI peptide from the carrier
XX
PS   Claim 10; Page 119; 186pp; English.
XX
CC   A new recombinant DNA vector construct has been developed which encodes
CC   a fusion protein and is suitable for introduction into a bacterial host.
CC   The vector comprises: (a) DNA encoding at least one cationic
CC   bactericidal/permeability increasing peptide (BPI), (b) DNA encoding a
CC   carrier protein, and (c) DNA encoding an amino acid (aa) cleavage site
CC   located between (a) and (b). The present sequence represents a
CC   specifically claimed BPI peptide. The peptides have many uses including
CC   the treatment of bacterial and fungal infections. BPI peptides also
CC   bind to endotoxins and heparin, neutralising their effects. The
CC   peptides have further been shown to inhibit angiogenesis (partly due to
CC   heparin-binding activity). The fusion proteins have been found to be
CC   expressed in large amounts without significant proteolysis, and in some
CC   cases are actually secreted from the host cells. This allows the
CC   indirect production of anti-microbial BPI peptides in microbial hosts.
XX
SQ   Sequence   7 AA;

  Query Match             100.0%;  Score 31;  DB 18;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 9.3e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ||||||
Db          2 LIQLFH 7


RESULT 10
AAW43732
ID   AAW43732 standard; peptide; 7 AA.
XX
AC   AAW43732;
XX
DT   20-APR-1998  (first entry)
XX
DE   Bactericidal/permeability increasing peptide XMP.310.
XX
KW   Bactericidal/permeability increasing peptide; BPI; fusion protein;
KW   bacterial infection; fungal infection; endotoxin; heparin;
KW   angiogenesis; fungicidal; recombinant DNA; vector.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   7
FT                   /note= "Amidated"
XX
PN   WO9735009-A1.
XX
PD   25-SEP-1997.
XX
PF   18-MAR-1997;   97WO-US05287.
XX
PR   22-MAR-1996;   96US-0621803.
XX
PA   (XOMA ) XOMA CORP.
XX
PI   Better MD;
XX
DR   WPI; 1997-480215/44.
XX
PT   Recombinant production of bactericidal/permeability increasing
PT   protein - by expression as a fusion protein in microbial host cells,
PT   then cleaving the BPI peptide from the carrier
XX
PS   Claim 10; Page 119; 186pp; English.
XX
CC   A new recombinant DNA vector construct has been developed which encodes
CC   a fusion protein and is suitable for introduction into a bacterial host.
CC   The vector comprises: (a) DNA encoding at least one cationic
CC   bactericidal/permeability increasing peptide (BPI), (b) DNA encoding a
CC   carrier protein, and (c) DNA encoding an amino acid (aa) cleavage site
CC   located between (a) and (b). The present sequence represents a
```

```
CC   specifically claimed BPI peptide. The peptides have many uses including
CC   the treatment of bacterial and fungal infections. BPI peptides also
CC   bind to endotoxins and heparin, neutralising their effects. The
CC   peptides have further been shown to inhibit angiogenesis (partly due to
CC   heparin-binding activity). The fusion proteins have been found to be
CC   expressed in large amounts without significant proteolysis, and in some
CC   cases are actually secreted from the host cells. This allows the
CC   indirect production of anti-microbial BPI peptides in microbial hosts.
XX
SQ   Sequence   7 AA;

  Query Match             100.0%;  Score 31;  DB 16;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 9.3e-05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ||||||
Db          1 LIQLFH 6


RESULT 11
AAY00519
ID   AAY00519 standard; Peptide; 7 AA.
XX
AC   AAY00519;
XX
DT   07-MAY-1999  (first entry)
XX
DE   Antifungal peptide XMP.310.
XX
KW   Antifungal; BPI; bactericidal/permeability increasing protein;
KW   Candida infection.
XX
OS   Synthetic.
XX
PN   US5858974-A.
XX
PD   12-JAN-1999.
XX
PF   21-MAR-1996;   96US-0621259.
XX
PR   21-MAR-1996;   96US-0621259.
PR   20-JUL-1995;   95US-0504841.
XX
PA   (XOMA ) XOMA CORP.
XX
PI   Fadem MB,  Lim E,  Little RG;
XX
DR   WPI; 1999-119956/10.
XX
PT   Antifungal peptides - comprising part of bactericidal or
PT   permeability-increasing protein sequence or related sequence
XX
PS   Disclosure; Columns 153-154; 132pp; English.
XX
CC   New peptides are provided which are based on Domain III (amino acids
CC   142-169) of human bactericidal/permeability-increasing protein (BPI).
CC   The peptides all have a C-terminal amide. More particularly, the Claims
CC   relate to: (1) a peptide that has an amino acid sequence of human BPI
CC   from position 148 to position 161 (KSKVGWLIQLFHKK) and variants of the
CC   sequence having antifungal activity; and (2) an antifungal peptide
CC   having 6-14 amino acids comprising (a) a core sequence selected from
CC   LIQL, IQLF, WLIQL, LIQLF and WLIQLF and (b) one or more cationic amino
CC   acids selected from K, R, H, Orn (ornithine) and Dab (diaminobutyric
CC   acid) at the N and/or C terminus of the core sequence. The new peptides
CC   are used for killing or inhibiting replication of fungi in vitro; and
CC   for treating fungal infections in vivo, in particular infections of
CC   Candida, Aspergillus or Cryptococcus spp., especially C. albicans, C.
CC   krusei, C. lusitaniae, C. parapsilosis or C. tropicalis. The peptide
CC   can be administered topically, intravenously, orally or as an aerosol,
CC   optionally together with a non-peptide antifungal agent.
XX
SQ   Sequence   7 AA;

  Query Match             100.0%;  Score 31;  DB 20;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 9.3e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ||||||
Db          1 LIQLFH 6


RESULT 12
AAY00518
ID   AAY00518 standard; Peptide; 7 AA.
XX
AC   AAY00518;
XX
DT   07-MAY-1999  (first entry)
XX
DE   Antifungal peptide XMP.309.
XX
KW   Antifungal; BPI; bactericidal/permeability increasing protein;
KW   Candida infection.
XX
OS   Synthetic.
XX
PN   US5858974-A.
XX
PD   12-JAN-1999.
XX
PF   21-MAR-1996;   96US-0621259.
XX
PR   21-MAR-1996;   96US-0621259.
PR   20-JUL-1995;   95US-0504841.
XX
PA   (XOMA ) XOMA CORP.
XX
PI   Fadem MB,  Lim E,  Little RG;
XX
DR   WPI; 1999-119956/10.
XX
PT   Antifungal peptides - comprising part of bactericidal or
PT   permeability-increasing protein sequence or related sequence
XX
PS   Disclosure; Columns 153-154; 132pp; English.
XX
CC   New peptides are provided which are based on Domain III (amino acids
CC   142-169) of human bactericidal/permeability-increasing protein (BPI).
CC   The peptides all have a C-terminal amide. More particularly, the Claims
CC   relate to: (1) a peptide that has an amino acid sequence of human BPI
CC   from position 148 to position 161 (KSKVGWLIQLFHKK) and variants of the
CC   sequence having antifungal activity; and (2) an antifungal peptide
CC   having 6-14 amino acids comprising (a) a core sequence selected from
CC   LIQL, IQLF, WLIQL, LIQLF and WLIQLF and (b) one or more cationic amino
CC   acids selected from K, R, H, Orn (ornithine) and Dab (diaminobutyric
CC   acid) at the N and/or C terminus of the core sequence. The new peptides
CC   are used for killing or inhibiting replication of fungi in vitro; and
CC   for treating fungal infections in vivo, in particular infections of
CC   Candida, Aspergillus or Cryptococcus spp., especially C. albicans, C.
CC   krusei, C. lusitaniae, C. parapsilosis or C. tropicalis. The peptide
CC   can be administered topically, intravenously, orally or as an aerosol,
CC   optionally together with a non-peptide antifungal agent.
XX
SQ   Sequence   7 AA;

  Query Match             100.0%;  Score 31;  DB 20;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 9.3e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ||||||
Db          2 LIQLFH 7
```

```
RESULT 13
AAB65442
ID   AAB65442 standard; Peptide; 7 AA.
XX
AC   AAB65442;
XX
DT   27-MAR-2001  (first entry)
XX
DE   Anti-fungal peptide XMP.309.
XX
KW   Human; BPI; antifungal; polymorphonuclear leukocyte; neutrophil;
KW   bactericidal/permeability-increasing protein; bactericidal;
KW   fungal infection.
XX
OS   Homo sapiens.
XX
PN   US6156730-A.
XX
PD   05-DEC-2000.
XX
PF   08-JAN-1999;   99US-0227659.
XX
PR   21-MAR-1996;   96US-0621259.
PR   12-MAR-1993;   93US-0030644.
PR   15-JUL-1993;   93US-0093202.
PR   14-JAN-1994;   94US-0183222.
PR   11-MAR-1994;   94US-0209762.
PR   11-JUL-1994;   94US-0273540.
PR   15-SEP-1994;   94US-0306473.
PR   13-JAN-1995;   95US-0372105.
PR   20-JUL-1995;   95US-0504841.
XX
PA   (XOMA ) XOMA CORP.
XX
PI   Lim E,  Fadem MB,  Little RG;
XX
DR   WPI; 2001-090160/10.
XX
PT   Novel anti-fungal peptides derived from domain III of
PT   bactericidal/permeability-increasing protein useful for killing or
PT   inhibiting replication of fungi and for treating fungal infections  -
XX
PS   Example 2; Columns 157-158; 134pp; English.
XX
CC   The present invention relates to antifungal peptides (see
CC   AAB65301-B65550) derived from Domain III (amino acids 142-169) of
CC   bactericidal/permeability-increasing protein (BPI). The present sequence
CC   is one such antifungal peptide. BPI is a protein isolated from the
CC   granules of mammalian polymorphonuclear leukocytes (PMNs or
CC   neutrophils). BPI has potent bactericidal activity against a broad range
CC   of gram-negative bacteria. The peptides of the present invention are
CC   useful for killing or inhibiting replication of fungi, and treating
CC   infections caused by fungus belonging to Candida, Aspergillus,
CC   Cryptococcus species such as C.albicans, C.glabrata, C.krusei,
CC   C.lusitaniae, C.parapsilosis and C.tropicalis.
XX
SQ   Sequence   7 AA;

  Query Match             100.0%;  Score 31;  DB 22;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 9.3e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ||||||
Db          2 LIQLFH 7


RESULT 14
AAB65443
ID   AAB65443 standard; Peptide; 7 AA.
XX
AC   AAB65443;
XX
DT   27-MAR-2001  (first entry)
XX
DE   Anti-fungal peptide XMP.310.
XX
KW   Human; BPI; antifungal; polymorphonuclear leukocyte; neutrophil;
KW   bactericidal/permeability-increasing protein; bactericidal;
KW   fungal infection.
XX
OS   Homo sapiens.
XX
PN   US6156730-A.
XX
PD   05-DEC-2000.
XX
PF   08-JAN-1999;   99US-0227659.
XX
PR   21-MAR-1996;   96US-0621259.
PR   12-MAR-1993;   93US-0030644.
PR   15-JUL-1993;   93US-0093202.
PR   14-JAN-1994;   94US-0183222.
PR   11-MAR-1994;   94US-0209762.
PR   11-JUL-1994;   94US-0273540.
PR   15-SEP-1994;   94US-0306473.
PR   13-JAN-1995;   95US-0372105.
PR   20-JUL-1995;   95US-0504841.
XX
PA   (XOMA ) XOMA CORP.
XX
PI   Lim E,  Fadem MB,  Little RG;
XX
DR   WPI; 2001-090160/10.
XX
PT   Novel anti-fungal peptides derived from domain III of
PT   bactericidal/permeability-increasing protein useful for killing or
PT   inhibiting replication of fungi and for treating fungal infections  -
XX
PS   Example 2; Columns 157-158; 134pp; English.
XX
CC   The present invention relates to antifungal peptides (see
CC   AAB65301-B65550) derived from Domain III (amino acids 142-169) of
CC   bactericidal/permeability-increasing protein (BPI). The present sequence
CC   is one such antifungal peptide. BPI is a protein isolated from the
CC   granules of mammalian polymorphonuclear leukocytes (PMNs or
CC   neutrophils). BPI has potent bactericidal activity against a broad range
CC   of gram-negative bacteria. The peptides of the present invention are
CC   useful for killing or inhibiting replication of fungi, and treating
CC   infections caused by fungus belonging to Candida, Aspergillus,
CC   Cryptococcus species such as C.albicans, C.glabrata, C.krusei,
CC   C.lusitaniae, C.parapsilosis and C.tropicalis.
XX
SQ   Sequence   7 AA;

  Query Match             100.0%;  Score 31;  DB 22;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 9.3e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ||||||
Db          1 LIQLFH 6


RESULT 15
AAW04017
ID   AAW04017 standard; peptide; 8 AA.
XX
AC   AAW04017;
XX
DT   31-OCT-1996  (first entry)
XX
DE   Antifungal peptide XMP.302.
XX
KW   Antifungal peptide; inhibitor; Domain III; polymorphonuclear leukocyte;
KW   bactericidal/permeability-increasing protein; BPI; mammalian; PMN; fungi;
```

KW neutrophil; replication inhibitor; fungal infection; Aspergillus;
KW Cryptococcus; Candida; C.albicans; C.galabrat; C.krusei; C.lusitaniae;
KW C.parapsilosis; C.tropicalis; therapy.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 8
FT /note= "amidated"
XX
PN WO9608509-A1.
XX
PD 21-MAR-1996.
XX
PF 20-JUL-1995; 95WO-US09262.
XX
PR 13-JAN-1995; 95US-0372105.
PR 15-SEP-1994; 94US-0306473.
XX
PA (XOMA) XOMA CORP.
XX
PI Fadem MB, Lim E, Little RG;
XX
DR WPI; 1996-179900/18.
XX
PT Antifungal peptide(s) derived from Domain III of BPI protein - used
PT in vitro for killing or inhibiting replication of fungi, esp.
PT Candida species
XX
PS Claim 5; Page 144; 199pp; English.
XX
CC AAW04000-W04160 represent antifungal peptides. These sequences are
CC based on (or derived from) Domain III of the
CC bactericidal/permeability-increasing protein (BPI). BPI is a protein
CC that can be isolated from the granules of mammalian polymorphonuclear
CC leukocytes (PMNs or neutrophils). These antifungal peptides can be used
CC for killing, or inhibiting replication of, fungi in vitro. These
CC sequences can also be used for treatment of a fungal infection involving
CC fungi from the species Candida, Aspergillus and Cryptococcus. The
CC sequences are especially useful for treating C.albicans, C.galabrat,
CC C.krusei, C.lusitaniae, C.parapsilosis and C.tropicalis infections.
XX
SQ Sequence 8 AA;

Query Match 100.0%; Score 31; DB 17; Length 8;
Best Local Similarity 100.0%; Pred. No. 9.3e-05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LIQLFH 6
||||||
Db 3 LIQLFH 8

Search completed: October 14, 2003, 14:41:30
Job time : 25.9167 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 14:33:23 ; Search time 8.5 Seconds
(without alignments)
29.866 Million cell updates/sec

Title: US-09-881-490-150
Perfect score: 31
Sequence: 1 LIQLFH 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	31	100.0	6	2	US-08-621-259A-150	Sequence 150, App
2	31	100.0	6	5	PCT-US95-09262-150	Sequence 150, App
3	31	100.0	7	2	US-08-621-803-175	Sequence 175, App
4	31	100.0	7	2	US-08-621-803-176	Sequence 176, App
5	31	100.0	7	2	US-08-621-259A-142	Sequence 142, App
6	31	100.0	7	2	US-08-621-259A-143	Sequence 143, App
7	31	100.0	7	3	US-09-217-352-175	Sequence 175, App
8	31	100.0	7	3	US-09-217-352-176	Sequence 176, App
9	31	100.0	7	5	PCT-US95-09262-142	Sequence 142, App
10	31	100.0	7	5	PCT-US95-09262-143	Sequence 143, App
11	31	100.0	8	2	US-08-621-803-168	Sequence 168, App
12	31	100.0	8	2	US-08-621-803-169	Sequence 169, App
13	31	100.0	8	2	US-08-621-803-170	Sequence 170, App
14	31	100.0	8	2	US-08-621-259A-135	Sequence 135, App
15	31	100.0	8	2	US-08-621-259A-136	Sequence 136, App
16	31	100.0	8	2	US-08-621-259A-137	Sequence 137, App
17	31	100.0	8	3	US-09-217-352-168	Sequence 168, App
18	31	100.0	8	3	US-09-217-352-169	Sequence 169, App
19	31	100.0	8	3	US-09-217-352-170	Sequence 170, App
20	31	100.0	8	5	PCT-US95-09262-135	Sequence 135, App
21	31	100.0	8	5	PCT-US95-09262-136	Sequence 136, App
22	31	100.0	8	5	PCT-US95-09262-137	Sequence 137, App
23	31	100.0	9	2	US-08-621-803-162	Sequence 162, App
24	31	100.0	9	2	US-08-621-803-163	Sequence 163, App
25	31	100.0	9	2	US-08-621-803-164	Sequence 164, App
26	31	100.0	9	2	US-08-621-259A-129	Sequence 129, App
27	31	100.0	9	2	US-08-621-259A-130	Sequence 130, App
28	31	100.0	9	2	US-08-621-259A-131	Sequence 131, App
29	31	100.0	9	3	US-09-217-352-162	Sequence 162, App
30	31	100.0	9	3	US-09-217-352-163	Sequence 163, App
31	31	100.0	9	3	US-09-217-352-164	Sequence 164, App
32	31	100.0	9	4	US-09-344-541A-10	Sequence 10, Appl
33	31	100.0	9	5	PCT-US95-09262-129	Sequence 129, App
34	31	100.0	9	5	PCT-US95-09262-130	Sequence 130, App
35	31	100.0	9	5	PCT-US95-09262-131	Sequence 131, App
36	31	100.0	10	2	US-08-621-803-157	Sequence 157, App
37	31	100.0	10	2	US-08-621-803-158	Sequence 158, App
38	31	100.0	10	2	US-08-621-803-159	Sequence 159, App
39	31	100.0	10	2	US-08-621-803-209	Sequence 209, App
40	31	100.0	10	2	US-08-621-803-210	Sequence 210, App
41	31	100.0	10	2	US-08-621-803-211	Sequence 211, App
42	31	100.0	10	2	US-08-621-803-212	Sequence 212, App
43	31	100.0	10	2	US-08-621-803-213	Sequence 213, App
44	31	100.0	10	2	US-08-621-803-214	Sequence 214, App
45	31	100.0	10	2	US-08-621-803-215	Sequence 215, App

ALIGNMENTS

```
RESULT 1
US-08-621-259A-150
; Sequence 150, Application US/08621259A
; Patent No. 5858974
;  GENERAL INFORMATION:
;    APPLICANT:  Little II, Roger G
;    APPLICANT:  Lim, Edward
;    APPLICANT:  Fadem, Mitchell B.
;    TITLE OF INVENTION:  Anti-Fungal Peptides
;    NUMBER OF SEQUENCES:  252
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McAndrews, Held & Malloy, Ltd.
;      STREET:  500 West Madison Street
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60661
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,259A
;      FILING DATE:  21-MAR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/504,841
;      FILING DATE:  20-JUL-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McNicholas, Janet M.
;      REGISTRATION NUMBER:  32,918
;      REFERENCE/DOCKET NUMBER:  11021US02
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/707-8889
;      TELEFAX:  312/707-9155
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  150:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.317"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
```

```
US-08-621-259A-150

  Query Match             100.0%;  Score 31;  DB 2;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ||||||
Db          1 LIQLFH 6


RESULT 2
PCT-US95-09262-150
; Sequence 150, Application PC/TUS9509262
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  Anti-Fungal Peptides
;    NUMBER OF SEQUENCES:  206
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/09262
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/372,105
;      FILING DATE:  13-JAN-95
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/306,473
;      FILING DATE:  15-SEP-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/273,540
;      FILING DATE:  11-JUL-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/209,762
;      FILING DATE:  11-MAR-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/183,222
;      FILING DATE:  14-JAN-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/093,202
;      FILING DATE:  15-JUL-93
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/030,644
;      FILING DATE:  12-MAR-93
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25,447
;      REFERENCE/DOCKET NUMBER:  27129/10040
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  150:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.317"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated"
PCT-US95-09262-150

  Query Match             100.0%;  Score 31;  DB 5;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 2.5e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ||||||
Db          1 LIQLFH 6


RESULT 3
US-08-621-803-175
; Sequence 175, Application US/08621803
; Patent No. 5851802
;  GENERAL INFORMATION:
;    APPLICANT:  Better, Marc D.
;    TITLE OF INVENTION:  Methods for Recombinant Microbial Production of
;    TITLE OF INVENTION:  Fusion Proteins and BPI-Derived Peptides
;    NUMBER OF SEQUENCES:  265
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,803
;      FILING DATE:  22-MAR-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25,447
;      REFERENCE/DOCKET NUMBER:  27129/33199
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  175:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.309"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-08-621-803-175

  Query Match             100.0%;  Score 31;  DB 2;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 2.5e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ||||||
Db          2 LIQLFH 7


RESULT 4
US-08-621-803-176
```

```
; Sequence 176, Application US/08621803
; Patent No. 5851802
;  GENERAL INFORMATION:
;    APPLICANT:  Better, Marc D.
;    TITLE OF INVENTION:  Methods for Recombinant Microbial Production of
;    TITLE OF INVENTION:  Fusion Proteins and BPI-Derived Peptides
;    NUMBER OF SEQUENCES:  265
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,803
;      FILING DATE:  22-MAR-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25,447
;      REFERENCE/DOCKET NUMBER:  27129/33199
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  176:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.310"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-08-621-803-176

  Query Match             100.0%;  Score 31;  DB 2;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ||||||
Db          1 LIQLFH 6


RESULT 5
US-08-621-259A-142
; Sequence 142, Application US/08621259A
; Patent No. 5858974
;  GENERAL INFORMATION:
;    APPLICANT:  Little II, Roger G
;    APPLICANT:  Lim, Edward
;    APPLICANT:  Fadem, Mitchell B.
;    TITLE OF INVENTION:  Anti-Fungal Peptides
;    NUMBER OF SEQUENCES:  252
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McAndrews, Held & Malloy, Ltd.
;      STREET:  500 West Madison Street
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60661
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,259A
;      FILING DATE:  21-MAR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/504,841
;      FILING DATE:  20-JUL-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McNicholas, Janet M.
;      REGISTRATION NUMBER:  32,918
;      REFERENCE/DOCKET NUMBER:  11021US02
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/707-8889
;      TELEFAX:  312/707-9155
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  142:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.309"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-08-621-259A-142

  Query Match             100.0%;  Score 31;  DB 2;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ||||||
Db          2 LIQLFH 7


RESULT 6
US-08-621-259A-143
; Sequence 143, Application US/08621259A
; Patent No. 5858974
;  GENERAL INFORMATION:
;    APPLICANT:  Little II, Roger G
;    APPLICANT:  Lim, Edward
;    APPLICANT:  Fadem, Mitchell B.
;    TITLE OF INVENTION:  Anti-Fungal Peptides
;    NUMBER OF SEQUENCES:  252
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McAndrews, Held & Malloy, Ltd.
;      STREET:  500 West Madison Street
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60661
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,259A
;      FILING DATE:  21-MAR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/504,841
;      FILING DATE:  20-JUL-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McNicholas, Janet M.
```

```
;       REGISTRATION NUMBER:  32,918
;       REFERENCE/DOCKET NUMBER:  11021US02
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  312/707-8889
;       TELEFAX:  312/707-9155
;       TELEX:
;   INFORMATION FOR SEQ ID NO:  143:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  7 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
;     FEATURE:
;       NAME/KEY:  misc_feature
;       OTHER INFORMATION:  "XMP.310"
;     FEATURE:
;       NAME/KEY:  Modified-site
;       LOCATION:  C-Terminus
;       OTHER INFORMATION:  /label= Amidation
;       OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-08-621-259A-143

  Query Match             100.0%;  Score 31;  DB 2;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ||||||
Db          1 LIQLFH 6


RESULT 7
US-09-217-352-175
; Sequence 175, Application US/09217352
; Patent No. 6274344
;  GENERAL INFORMATION:
;    APPLICANT:  Better, Marc D.
;    TITLE OF INVENTION:  Methods for Recombinant Microbial Production of
;    TITLE OF INVENTION:  Fusion Proteins and BPI-Derived Peptides
;    NUMBER OF SEQUENCES:  265
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/217,352
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/621,803
;      FILING DATE:  22-MAR-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25,447
;      REFERENCE/DOCKET NUMBER:  27129/33199
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  175:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.309"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-09-217-352-175

  Query Match             100.0%;  Score 31;  DB 3;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 2.5e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ||||||
Db          2 LIQLFH 7


RESULT 8
US-09-217-352-176
; Sequence 176, Application US/09217352
; Patent No. 6274344
;  GENERAL INFORMATION:
;    APPLICANT:  Better, Marc D.
;    TITLE OF INVENTION:  Methods for Recombinant Microbial Production of
;    TITLE OF INVENTION:  Fusion Proteins and BPI-Derived Peptides
;    NUMBER OF SEQUENCES:  265
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/217,352
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/621,803
;      FILING DATE:  22-MAR-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25,447
;      REFERENCE/DOCKET NUMBER:  27129/33199
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  176:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.310"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-09-217-352-176

  Query Match             100.0%;  Score 31;  DB 3;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 2.5e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 LIQLFH 6
              ||||||
Db          1 LIQLFH 6


RESULT 9
PCT-US95-09262-142
; Sequence 142, Application PC/TUS9509262
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  Anti-Fungal Peptides
;    NUMBER OF SEQUENCES:  206
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/09262
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/372,105
;      FILING DATE:  13-JAN-95
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/306,473
;      FILING DATE:  15-SEP-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/273,540
;      FILING DATE:  11-JUL-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/209,762
;      FILING DATE:  11-MAR-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/183,222
;      FILING DATE:  14-JAN-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/093,202
;      FILING DATE:  15-JUL-93
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/030,644
;      FILING DATE:  12-MAR-93
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25,447
;      REFERENCE/DOCKET NUMBER:  27129/10040
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  142:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.309"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated"
PCT-US95-09262-142

  Query Match             100.0%;  Score 31;  DB 5;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ||||||
Db          2 LIQLFH 7


RESULT 10
PCT-US95-09262-143
; Sequence 143, Application PC/TUS9509262
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  Anti-Fungal Peptides
;    NUMBER OF SEQUENCES:  206
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/09262
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/372,105
;      FILING DATE:  13-JAN-95
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/306,473
;      FILING DATE:  15-SEP-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/273,540
;      FILING DATE:  11-JUL-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/209,762
;      FILING DATE:  11-MAR-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/183,222
;      FILING DATE:  14-JAN-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/093,202
;      FILING DATE:  15-JUL-93
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/030,644
;      FILING DATE:  12-MAR-93
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25,447
;      REFERENCE/DOCKET NUMBER:  27129/10040
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  143:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.310"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated"
```

```
PCT-US95-09262-143

  Query Match             100.0%;  Score 31;  DB 5;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ||||||
Db          1 LIQLFH 6


RESULT 11
US-08-621-803-168
; Sequence 168, Application US/08621803
; Patent No. 5851802
;  GENERAL INFORMATION:
;    APPLICANT:  Better, Marc D.
;    TITLE OF INVENTION:  Methods for Recombinant Microbial Production of
;    TITLE OF INVENTION:  Fusion Proteins and BPI-Derived Peptides
;    NUMBER OF SEQUENCES:  265
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,803
;      FILING DATE:  22-MAR-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25,447
;      REFERENCE/DOCKET NUMBER:  27129/33199
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  168:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  8 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.302"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-08-621-803-168

  Query Match             100.0%;  Score 31;  DB 2;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ||||||
Db          3 LIQLFH 8


RESULT 12
US-08-621-803-169
; Sequence 169, Application US/08621803
; Patent No. 5851802
;  GENERAL INFORMATION:
;    APPLICANT:  Better, Marc D.
;    TITLE OF INVENTION:  Methods for Recombinant Microbial Production of
;    TITLE OF INVENTION:  Fusion Proteins and BPI-Derived Peptides
;    NUMBER OF SEQUENCES:  265
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,803
;      FILING DATE:  22-MAR-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25,447
;      REFERENCE/DOCKET NUMBER:  27129/33199
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  169:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  8 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.303"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-08-621-803-169

  Query Match             100.0%;  Score 31;  DB 2;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ||||||
Db          2 LIQLFH 7


RESULT 13
US-08-621-803-170
; Sequence 170, Application US/08621803
; Patent No. 5851802
;  GENERAL INFORMATION:
;    APPLICANT:  Better, Marc D.
;    TITLE OF INVENTION:  Methods for Recombinant Microbial Production of
;    TITLE OF INVENTION:  Fusion Proteins and BPI-Derived Peptides
;    NUMBER OF SEQUENCES:  265
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
```

```
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,803
;      FILING DATE:  22-MAR-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25,447
;      REFERENCE/DOCKET NUMBER:  27129/33199
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  170:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  8 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.304"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-08-621-803-170

  Query Match             100.0%;  Score 31;  DB 2;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 2.5e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ||||||
Db          1 LIQLFH 6


RESULT 14
US-08-621-259A-135
; Sequence 135, Application US/08621259A
; Patent No. 5858974
;  GENERAL INFORMATION:
;    APPLICANT:  Little II, Roger G
;    APPLICANT:  Lim, Edward
;    APPLICANT:  Fadem, Mitchell B.
;    TITLE OF INVENTION:  Anti-Fungal Peptides
;    NUMBER OF SEQUENCES:  252
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McAndrews, Held & Malloy, Ltd.
;      STREET:  500 West Madison Street
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60661
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,259A
;      FILING DATE:  21-MAR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/504,841
;      FILING DATE:  20-JUL-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McNicholas, Janet M.
;      REGISTRATION NUMBER:  32,918
;      REFERENCE/DOCKET NUMBER:  11021US02
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/707-8889
;      TELEFAX:  312/707-9155
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  135:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  8 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.302"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-08-621-259A-135

  Query Match             100.0%;  Score 31;  DB 2;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 2.5e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ||||||
Db          3 LIQLFH 8


RESULT 15
US-08-621-259A-136
; Sequence 136, Application US/08621259A
; Patent No. 5858974
;  GENERAL INFORMATION:
;    APPLICANT:  Little II, Roger G
;    APPLICANT:  Lim, Edward
;    APPLICANT:  Fadem, Mitchell B.
;    TITLE OF INVENTION:  Anti-Fungal Peptides
;    NUMBER OF SEQUENCES:  252
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McAndrews, Held & Malloy, Ltd.
;      STREET:  500 West Madison Street
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60661
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,259A
;      FILING DATE:  21-MAR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/504,841
;      FILING DATE:  20-JUL-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McNicholas, Janet M.
;      REGISTRATION NUMBER:  32,918
;      REFERENCE/DOCKET NUMBER:  11021US02
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/707-8889
;      TELEFAX:  312/707-9155
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  136:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  8 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.303"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
```

; OTHER INFORMATION: /note= "The C-Terminus is Amidated."
US-08-621-259A-136

Query Match 100.0%; Score 31; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 LIQLFH 6
              ||||||
Db          2 LIQLFH 7
```

Search completed: October 14, 2003, 14:42:45
Job time : 9.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 14:34:14 ; Search time 14.625 Seconds
(without alignments)
66.104 Million cell updates/sec

Title: US-09-881-490-150
Perfect score: 31
Sequence: 1 LIQLFH 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 600653 seqs, 161128416 residues

Total number of hits satisfying chosen parameters: 600653

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	31	100.0	6	9	US-09-881-490-150	Sequence 150, App
2	31	100.0	7	9	US-09-765-527-175	Sequence 175, App
3	31	100.0	7	9	US-09-765-527-176	Sequence 176, App
4	31	100.0	7	9	US-09-881-490-142	Sequence 142, App
5	31	100.0	7	9	US-09-881-490-143	Sequence 143, App
6	31	100.0	8	9	US-09-765-527-168	Sequence 168, App
7	31	100.0	8	9	US-09-765-527-169	Sequence 169, App
8	31	100.0	8	9	US-09-765-527-170	Sequence 170, App
9	31	100.0	8	9	US-09-881-490-135	Sequence 135, App
10	31	100.0	8	9	US-09-881-490-136	Sequence 136, App
11	31	100.0	8	9	US-09-881-490-137	Sequence 137, App
12	31	100.0	9	9	US-09-765-527-162	Sequence 162, App
13	31	100.0	9	9	US-09-765-527-163	Sequence 163, App
14	31	100.0	9	9	US-09-765-527-164	Sequence 164, App
15	31	100.0	9	9	US-09-881-490-129	Sequence 129, App
16	31	100.0	9	9	US-09-881-490-130	Sequence 130, App
17	31	100.0	9	9	US-09-881-490-131	Sequence 131, App
18	31	100.0	10	9	US-09-765-527-157	Sequence 157, App
19	31	100.0	10	9	US-09-765-527-158	Sequence 158, App
20	31	100.0	10	9	US-09-765-527-159	Sequence 159, App
21	31	100.0	10	9	US-09-765-527-209	Sequence 209, App
22	31	100.0	10	9	US-09-765-527-210	Sequence 210, App
23	31	100.0	10	9	US-09-765-527-211	Sequence 211, App
24	31	100.0	10	9	US-09-765-527-212	Sequence 212, App
25	31	100.0	10	9	US-09-765-527-213	Sequence 213, App
26	31	100.0	10	9	US-09-765-527-214	Sequence 214, App
27	31	100.0	10	9	US-09-765-527-215	Sequence 215, App
28	31	100.0	10	9	US-09-765-527-227	Sequence 227, App
29	31	100.0	10	9	US-09-881-490-124	Sequence 124, App
30	31	100.0	10	9	US-09-881-490-125	Sequence 125, App
31	31	100.0	10	9	US-09-881-490-126	Sequence 126, App
32	31	100.0	10	9	US-09-881-490-184	Sequence 184, App
33	31	100.0	10	9	US-09-881-490-185	Sequence 185, App
34	31	100.0	10	9	US-09-881-490-186	Sequence 186, App
35	31	100.0	10	9	US-09-881-490-187	Sequence 187, App
36	31	100.0	10	9	US-09-881-490-188	Sequence 188, App
37	31	100.0	10	9	US-09-881-490-190	Sequence 190, App
38	31	100.0	10	9	US-09-881-490-194	Sequence 194, App
39	31	100.0	10	9	US-09-881-490-195	Sequence 195, App
40	31	100.0	10	9	US-09-881-490-196	Sequence 196, App
41	31	100.0	10	9	US-09-881-490-197	Sequence 197, App
42	31	100.0	10	9	US-09-881-490-204	Sequence 204, App
43	31	100.0	10	14	US-10-006-557-11	Sequence 11, Appl
44	31	100.0	10	15	US-10-146-136-3	Sequence 3, Appli
45	31	100.0	10	15	US-10-146-136-5	Sequence 5, Appli

ALIGNMENTS

```
RESULT 1
US-09-881-490-150
; Sequence 150, Application US/09881490
; Patent No. US20020077298A1
;   GENERAL INFORMATION:
;        APPLICANT: Little II, Roger G.
;                   Lim, Edward
;                   Fadem, Mitchell B.
;        TITLE OF INVENTION: Anti-Fungal Peptides
;        NUMBER OF SEQUENCES: 211
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: McAndrews, Held & Malloy, Ltd.
;             STREET: 500 West Madison Street, 34th FloorDrive
;             CITY: Chicago
;             STATE: Illinois
;             COUNTRY: United States of America
;             ZIP: 60661
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/881,490
;             FILING DATE: 14-Jun-2001
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/119,858
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: 08/372,105
;             FILING DATE: 13-JAN-95
;             APPLICATION NUMBER: 08/306,473
;             FILING DATE: 15-SEP-94
;             APPLICATION NUMBER: 08/273,540
;             FILING DATE: 11-JUL-94
;             APPLICATION NUMBER: 08/209,762
;             FILING DATE: 11-MAR-94
;             APPLICATION NUMBER: 08/183,222
;             FILING DATE: 14-JAN-94
```

```
;              APPLICATION NUMBER: 08/093,202
;              FILING DATE: 15-JUL-93
;              APPLICATION NUMBER: 08/030,644
;              FILING DATE: 12-MAR-93
;         ATTORNEY/AGENT INFORMATION:
;              NAME: McNicholas, Janet M.
;              REGISTRATION NUMBER: 32,918
;              REFERENCE/DOCKET NUMBER: 100-238/11021US01
;         TELECOMMUNICATION INFORMATION:
;              TELEPHONE: 312/707-8889
;              TELEFAX: 312/707-9155
;              TELEX: 650 388-1248
;    INFORMATION FOR SEQ ID NO: 150:
;         SEQUENCE CHARACTERISTICS:
;              LENGTH: 6 amino acids
;              TYPE: amino acid
;              TOPOLOGY: linear
;         MOLECULE TYPE: peptide
;         FEATURE:
;              NAME/KEY:  misc_feature
;              OTHER INFORMATION:  "XMP.317"
;         FEATURE:
;              NAME/KEY:  Modified-site
;              LOCATION:  C-Terminus
;              OTHER INFORMATION:  /label= Amidation
;              /note= "The C-Terminus is Amidated"
;         SEQUENCE DESCRIPTION: SEQ ID NO: 150:
US-09-881-490-150

  Query Match             100.0%;  Score 31;  DB 9;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 5.3e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ||||||
Db          1 LIQLFH 6


RESULT 2
US-09-765-527-175
; Sequence 175, Application US/09765527
; Patent No. US20020006638A1
;   GENERAL INFORMATION:
;         APPLICANT: Better, Marc D.
;         TITLE OF INVENTION: Methods for Recombinant Microbial Production of
;                             Fusion Proteins and BPI-Derived Peptides
;         NUMBER OF SEQUENCES: 265
;         CORRESPONDENCE ADDRESS:
;              ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
;              STREET: 6300 Sears Tower, 233 South Wacker Drive
;              CITY: Chicago
;              STATE: Illinois
;              COUNTRY: United States of America
;              ZIP: 60606-6402
;         COMPUTER READABLE FORM:
;              MEDIUM TYPE: Floppy disk
;              COMPUTER: IBM PC compatible
;              OPERATING SYSTEM: PC-DOS/MS-DOS
;              SOFTWARE: PatentIn Release #1.0, Version #1.25
;         CURRENT APPLICATION DATA:
;              APPLICATION NUMBER: US/09/765,527
;              FILING DATE: 18-Jan-2001
;         PRIOR APPLICATION DATA:
;              APPLICATION NUMBER: 08/621,803
;              FILING DATE: <Unknown>
;         ATTORNEY/AGENT INFORMATION:
;              NAME: Borun, Michael F.
;              REGISTRATION NUMBER: 25,447
;              REFERENCE/DOCKET NUMBER: 27129/33199
;         TELECOMMUNICATION INFORMATION:
;              TELEPHONE: 312/474-6300
;              TELEFAX: 312/474-0448
;              TELEX: 25-3856
;    INFORMATION FOR SEQ ID NO: 175:
;         SEQUENCE CHARACTERISTICS:
;              LENGTH: 7 amino acids
;              TYPE: amino acid
;              TOPOLOGY: linear
;         MOLECULE TYPE: peptide
;         FEATURE:
;              NAME/KEY:  misc_feature
;              OTHER INFORMATION:  "XMP.309"
;         FEATURE:
;              NAME/KEY:  Modified-site
;              LOCATION:  C-Terminus
;              OTHER INFORMATION:  /label= Amidation
;              /note= "The C-Terminus is Amidated."
;         SEQUENCE DESCRIPTION: SEQ ID NO: 175:
US-09-765-527-175

  Query Match             100.0%;  Score 31;  DB 9;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 5.3e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ||||||
Db          2 LIQLFH 7


RESULT 3
US-09-765-527-176
; Sequence 176, Application US/09765527
; Patent No. US20020006638A1
;   GENERAL INFORMATION:
;         APPLICANT: Better, Marc D.
;         TITLE OF INVENTION: Methods for Recombinant Microbial Production of
;                             Fusion Proteins and BPI-Derived Peptides
;         NUMBER OF SEQUENCES: 265
;         CORRESPONDENCE ADDRESS:
;              ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
;              STREET: 6300 Sears Tower, 233 South Wacker Drive
;              CITY: Chicago
;              STATE: Illinois
;              COUNTRY: United States of America
;              ZIP: 60606-6402
;         COMPUTER READABLE FORM:
;              MEDIUM TYPE: Floppy disk
;              COMPUTER: IBM PC compatible
;              OPERATING SYSTEM: PC-DOS/MS-DOS
;              SOFTWARE: PatentIn Release #1.0, Version #1.25
;         CURRENT APPLICATION DATA:
;              APPLICATION NUMBER: US/09/765,527
;              FILING DATE: 18-Jan-2001
;         PRIOR APPLICATION DATA:
;              APPLICATION NUMBER: 08/621,803
;              FILING DATE: <Unknown>
;         ATTORNEY/AGENT INFORMATION:
;              NAME: Borun, Michael F.
;              REGISTRATION NUMBER: 25,447
;              REFERENCE/DOCKET NUMBER: 27129/33199
;         TELECOMMUNICATION INFORMATION:
;              TELEPHONE: 312/474-6300
;              TELEFAX: 312/474-0448
;              TELEX: 25-3856
;    INFORMATION FOR SEQ ID NO: 176:
;         SEQUENCE CHARACTERISTICS:
;              LENGTH: 7 amino acids
;              TYPE: amino acid
;              TOPOLOGY: linear
;         MOLECULE TYPE: peptide
;         FEATURE:
;              NAME/KEY:  misc_feature
;              OTHER INFORMATION:  "XMP.310"
;         FEATURE:
;              NAME/KEY:  Modified-site
;              LOCATION:  C-Terminus
```

```
;             OTHER INFORMATION:  /label= Amidation
;             /note= "The C-Terminus is Amidated."
;        SEQUENCE DESCRIPTION: SEQ ID NO: 176:
US-09-765-527-176

  Query Match             100.0%;  Score 31;  DB 9;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ||||||
Db          1 LIQLFH 6


RESULT 4
US-09-881-490-142
; Sequence 142, Application US/09881490
; Patent No. US20020077298A1
;   GENERAL INFORMATION:
;        APPLICANT: Little II, Roger G.
;                   Lim, Edward
;                   Fadem, Mitchell B.
;        TITLE OF INVENTION: Anti-Fungal Peptides
;        NUMBER OF SEQUENCES: 211
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: McAndrews, Held & Malloy, Ltd.
;             STREET: 500 West Madison Street, 34th FloorDrive
;             CITY: Chicago
;             STATE: Illinois
;             COUNTRY: United States of America
;             ZIP: 60661
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/881,490
;             FILING DATE: 14-Jun-2001
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/119,858
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: 08/372,105
;             FILING DATE: 13-JAN-95
;             APPLICATION NUMBER: 08/306,473
;             FILING DATE: 15-SEP-94
;             APPLICATION NUMBER: 08/273,540
;             FILING DATE: 11-JUL-94
;             APPLICATION NUMBER: 08/209,762
;             FILING DATE: 11-MAR-94
;             APPLICATION NUMBER: 08/183,222
;             FILING DATE: 14-JAN-94
;             APPLICATION NUMBER: 08/093,202
;             FILING DATE: 15-JUL-93
;             APPLICATION NUMBER: 08/030,644
;             FILING DATE: 12-MAR-93
;        ATTORNEY/AGENT INFORMATION:
;             NAME: McNicholas, Janet M.
;             REGISTRATION NUMBER: 32,918
;             REFERENCE/DOCKET NUMBER: 100-238/11021US01
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 312/707-8889
;             TELEFAX: 312/707-9155
;             TELEX: 650 388-1248
;   INFORMATION FOR SEQ ID NO: 142:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 7 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        FEATURE:
;             NAME/KEY:  misc_feature
;             OTHER INFORMATION:  "XMP.309"
;        FEATURE:
;             NAME/KEY:  Modified-site
;             LOCATION:  C-Terminus
;             OTHER INFORMATION:  /label= Amidation
;             /note= "The C-Terminus is Amidated"
;        SEQUENCE DESCRIPTION: SEQ ID NO: 142:
US-09-881-490-142

  Query Match             100.0%;  Score 31;  DB 9;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ||||||
Db          2 LIQLFH 7


RESULT 5
US-09-881-490-143
; Sequence 143, Application US/09881490
; Patent No. US20020077298A1
;   GENERAL INFORMATION:
;        APPLICANT: Little II, Roger G.
;                   Lim, Edward
;                   Fadem, Mitchell B.
;        TITLE OF INVENTION: Anti-Fungal Peptides
;        NUMBER OF SEQUENCES: 211
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: McAndrews, Held & Malloy, Ltd.
;             STREET: 500 West Madison Street, 34th FloorDrive
;             CITY: Chicago
;             STATE: Illinois
;             COUNTRY: United States of America
;             ZIP: 60661
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/881,490
;             FILING DATE: 14-Jun-2001
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/119,858
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: 08/372,105
;             FILING DATE: 13-JAN-95
;             APPLICATION NUMBER: 08/306,473
;             FILING DATE: 15-SEP-94
;             APPLICATION NUMBER: 08/273,540
;             FILING DATE: 11-JUL-94
;             APPLICATION NUMBER: 08/209,762
;             FILING DATE: 11-MAR-94
;             APPLICATION NUMBER: 08/183,222
;             FILING DATE: 14-JAN-94
;             APPLICATION NUMBER: 08/093,202
;             FILING DATE: 15-JUL-93
;             APPLICATION NUMBER: 08/030,644
;             FILING DATE: 12-MAR-93
;        ATTORNEY/AGENT INFORMATION:
;             NAME: McNicholas, Janet M.
;             REGISTRATION NUMBER: 32,918
;             REFERENCE/DOCKET NUMBER: 100-238/11021US01
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 312/707-8889
;             TELEFAX: 312/707-9155
;             TELEX: 650 388-1248
;   INFORMATION FOR SEQ ID NO: 143:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 7 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
```

```
;       FEATURE:
;            NAME/KEY:  misc_feature
;            OTHER INFORMATION:  "XMP.310"
;       FEATURE:
;            NAME/KEY:  Modified-site
;            LOCATION:  C-Terminus
;            OTHER INFORMATION:  /label= Amidation
;            /note= "The C-Terminus is Amidated"
;       SEQUENCE DESCRIPTION: SEQ ID NO: 143:
US-09-881-490-143

  Query Match             100.0%;  Score 31;  DB 9;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ||||||
Db          1 LIQLFH 6


RESULT 6
US-09-765-527-168
; Sequence 168, Application US/09765527
; Patent No. US20020006638A1
;   GENERAL INFORMATION:
;       APPLICANT: Better, Marc D.
;       TITLE OF INVENTION: Methods for Recombinant Microbial Production of
;                           Fusion Proteins and BPI-Derived Peptides
;       NUMBER OF SEQUENCES: 265
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
;            STREET: 6300 Sears Tower, 233 South Wacker Drive
;            CITY: Chicago
;            STATE: Illinois
;            COUNTRY: United States of America
;            ZIP: 60606-6402
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/765,527
;            FILING DATE: 18-Jan-2001
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 08/621,803
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Borun, Michael F.
;            REGISTRATION NUMBER: 25,447
;            REFERENCE/DOCKET NUMBER: 27129/33199
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 312/474-6300
;            TELEFAX: 312/474-0448
;            TELEX: 25-3856
;   INFORMATION FOR SEQ ID NO: 168:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 8 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: peptide
;       FEATURE:
;            NAME/KEY:  misc_feature
;            OTHER INFORMATION:  "XMP.302"
;       FEATURE:
;            NAME/KEY:  Modified-site
;            LOCATION:  C-Terminus
;            OTHER INFORMATION:  /label= Amidation
;            /note= "The C-Terminus is Amidated."
;       SEQUENCE DESCRIPTION: SEQ ID NO: 168:
US-09-765-527-168

  Query Match             100.0%;  Score 31;  DB 9;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 5.3e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ||||||
Db          3 LIQLFH 8


RESULT 7
US-09-765-527-169
; Sequence 169, Application US/09765527
; Patent No. US20020006638A1
;   GENERAL INFORMATION:
;       APPLICANT: Better, Marc D.
;       TITLE OF INVENTION: Methods for Recombinant Microbial Production of
;                           Fusion Proteins and BPI-Derived Peptides
;       NUMBER OF SEQUENCES: 265
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
;            STREET: 6300 Sears Tower, 233 South Wacker Drive
;            CITY: Chicago
;            STATE: Illinois
;            COUNTRY: United States of America
;            ZIP: 60606-6402
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/765,527
;            FILING DATE: 18-Jan-2001
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 08/621,803
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Borun, Michael F.
;            REGISTRATION NUMBER: 25,447
;            REFERENCE/DOCKET NUMBER: 27129/33199
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 312/474-6300
;            TELEFAX: 312/474-0448
;            TELEX: 25-3856
;   INFORMATION FOR SEQ ID NO: 169:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 8 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: peptide
;       FEATURE:
;            NAME/KEY:  misc_feature
;            OTHER INFORMATION:  "XMP.303"
;       FEATURE:
;            NAME/KEY:  Modified-site
;            LOCATION:  C-Terminus
;            OTHER INFORMATION:  /label= Amidation
;            /note= "The C-Terminus is Amidated."
;       SEQUENCE DESCRIPTION: SEQ ID NO: 169:
US-09-765-527-169

  Query Match             100.0%;  Score 31;  DB 9;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 5.3e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ||||||
Db          2 LIQLFH 7


RESULT 8
US-09-765-527-170
; Sequence 170, Application US/09765527
; Patent No. US20020006638A1
```

```
;   GENERAL INFORMATION:
;        APPLICANT: Better, Marc D.
;        TITLE OF INVENTION: Methods for Recombinant Microbial Production of
;                            Fusion Proteins and BPI-Derived Peptides
;        NUMBER OF SEQUENCES: 265
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
;             STREET: 6300 Sears Tower, 233 South Wacker Drive
;             CITY: Chicago
;             STATE: Illinois
;             COUNTRY: United States of America
;             ZIP: 60606-6402
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/765,527
;             FILING DATE: 18-Jan-2001
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/621,803
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Borun, Michael F.
;             REGISTRATION NUMBER: 25,447
;             REFERENCE/DOCKET NUMBER: 27129/33199
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 312/474-6300
;             TELEFAX: 312/474-0448
;             TELEX: 25-3856
;   INFORMATION FOR SEQ ID NO: 170:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 8 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        FEATURE:
;             NAME/KEY:  misc_feature
;             OTHER INFORMATION:  "XMP.304"
;        FEATURE:
;             NAME/KEY:  Modified-site
;             LOCATION:  C-Terminus
;             OTHER INFORMATION:  /label= Amidation
;             /note= "The C-Terminus is Amidated."
;        SEQUENCE DESCRIPTION: SEQ ID NO: 170:
US-09-765-527-170

  Query Match             100.0%;  Score 31;  DB 9;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ||||||
Db          1 LIQLFH 6


RESULT 9
US-09-881-490-135
; Sequence 135, Application US/09881490
; Patent No. US20020077298A1
;   GENERAL INFORMATION:
;        APPLICANT: Little II, Roger G.
;                   Lim, Edward
;                   Fadem, Mitchell B.
;        TITLE OF INVENTION: Anti-Fungal Peptides
;        NUMBER OF SEQUENCES: 211
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: McAndrews, Held & Malloy, Ltd.
;             STREET: 500 West Madison Street, 34th FloorDrive
;             CITY: Chicago
;             STATE: Illinois
;             COUNTRY: United States of America
;             ZIP: 60661
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/881,490
;             FILING DATE: 14-Jun-2001
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/119,858
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: 08/372,105
;             FILING DATE: 13-JAN-95
;             APPLICATION NUMBER: 08/306,473
;             FILING DATE: 15-SEP-94
;             APPLICATION NUMBER: 08/273,540
;             FILING DATE: 11-JUL-94
;             APPLICATION NUMBER: 08/209,762
;             FILING DATE: 11-MAR-94
;             APPLICATION NUMBER: 08/183,222
;             FILING DATE: 14-JAN-94
;             APPLICATION NUMBER: 08/093,202
;             FILING DATE: 15-JUL-93
;             APPLICATION NUMBER: 08/030,644
;             FILING DATE: 12-MAR-93
;        ATTORNEY/AGENT INFORMATION:
;             NAME: McNicholas, Janet M.
;             REGISTRATION NUMBER: 32,918
;             REFERENCE/DOCKET NUMBER: 100-238/11021US01
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 312/707-8889
;             TELEFAX: 312/707-9155
;             TELEX: 650 388-1248
;   INFORMATION FOR SEQ ID NO: 135:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 8 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        FEATURE:
;             NAME/KEY:  misc_feature
;             OTHER INFORMATION:  "XMP.302"
;        FEATURE:
;             NAME/KEY:  Modified-site
;             LOCATION:  C-Terminus
;             OTHER INFORMATION:  /label= Amidation
;             /note= "The C-Terminus is Amidated"
;        SEQUENCE DESCRIPTION: SEQ ID NO: 135:
US-09-881-490-135

  Query Match             100.0%;  Score 31;  DB 9;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 5.3e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ||||||
Db          3 LIQLFH 8


RESULT 10
US-09-881-490-136
; Sequence 136, Application US/09881490
; Patent No. US20020077298A1
;   GENERAL INFORMATION:
;        APPLICANT: Little II, Roger G.
;                   Lim, Edward
;                   Fadem, Mitchell B.
;        TITLE OF INVENTION: Anti-Fungal Peptides
;        NUMBER OF SEQUENCES: 211
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: McAndrews, Held & Malloy, Ltd.
;             STREET: 500 West Madison Street, 34th FloorDrive
```

```
;             CITY: Chicago
;             STATE: Illinois
;             COUNTRY: United States of America
;             ZIP: 60661
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/881,490
;             FILING DATE: 14-Jun-2001
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/119,858
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: 08/372,105
;             FILING DATE: 13-JAN-95
;             APPLICATION NUMBER: 08/306,473
;             FILING DATE: 15-SEP-94
;             APPLICATION NUMBER: 08/273,540
;             FILING DATE: 11-JUL-94
;             APPLICATION NUMBER: 08/209,762
;             FILING DATE: 11-MAR-94
;             APPLICATION NUMBER: 08/183,222
;             FILING DATE: 14-JAN-94
;             APPLICATION NUMBER: 08/093,202
;             FILING DATE: 15-JUL-93
;             APPLICATION NUMBER: 08/030,644
;             FILING DATE: 12-MAR-93
;        ATTORNEY/AGENT INFORMATION:
;             NAME: McNicholas, Janet M.
;             REGISTRATION NUMBER: 32,918
;             REFERENCE/DOCKET NUMBER: 100-238/11021US01
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 312/707-8889
;             TELEFAX: 312/707-9155
;             TELEX: 650 388-1248
;    INFORMATION FOR SEQ ID NO: 136:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 8 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        FEATURE:
;             NAME/KEY:  misc_feature
;             OTHER INFORMATION:  "XMP.303"
;        FEATURE:
;             NAME/KEY:  Modified-site
;             LOCATION:  C-Terminus
;             OTHER INFORMATION:  /label= Amidation
;             /note= "The C-Terminus is Amidated"
;        SEQUENCE DESCRIPTION: SEQ ID NO: 136:
US-09-881-490-136

  Query Match             100.0%;  Score 31;  DB 9;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 5.3e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ||||||
Db          2 LIQLFH 7


RESULT 11
US-09-881-490-137
; Sequence 137, Application US/09881490
; Patent No. US20020077298A1
;   GENERAL INFORMATION:
;        APPLICANT: Little II, Roger G.
;                   Lim, Edward
;                   Fadem, Mitchell B.
;        TITLE OF INVENTION: Anti-Fungal Peptides
;        NUMBER OF SEQUENCES: 211
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: McAndrews, Held & Malloy, Ltd.
;             STREET: 500 West Madison Street, 34th FloorDrive
;             CITY: Chicago
;             STATE: Illinois
;             COUNTRY: United States of America
;             ZIP: 60661
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/881,490
;             FILING DATE: 14-Jun-2001
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/119,858
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: 08/372,105
;             FILING DATE: 13-JAN-95
;             APPLICATION NUMBER: 08/306,473
;             FILING DATE: 15-SEP-94
;             APPLICATION NUMBER: 08/273,540
;             FILING DATE: 11-JUL-94
;             APPLICATION NUMBER: 08/209,762
;             FILING DATE: 11-MAR-94
;             APPLICATION NUMBER: 08/183,222
;             FILING DATE: 14-JAN-94
;             APPLICATION NUMBER: 08/093,202
;             FILING DATE: 15-JUL-93
;             APPLICATION NUMBER: 08/030,644
;             FILING DATE: 12-MAR-93
;        ATTORNEY/AGENT INFORMATION:
;             NAME: McNicholas, Janet M.
;             REGISTRATION NUMBER: 32,918
;             REFERENCE/DOCKET NUMBER: 100-238/11021US01
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 312/707-8889
;             TELEFAX: 312/707-9155
;             TELEX: 650 388-1248
;    INFORMATION FOR SEQ ID NO: 137:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 8 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        FEATURE:
;             NAME/KEY:  misc_feature
;             OTHER INFORMATION:  "XMP.304"
;        FEATURE:
;             NAME/KEY:  Modified-site
;             LOCATION:  C-Terminus
;             OTHER INFORMATION:  /label= Amidation
;             /note= "The C-Terminus is Amidated"
;        SEQUENCE DESCRIPTION: SEQ ID NO: 137:
US-09-881-490-137

  Query Match             100.0%;  Score 31;  DB 9;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 5.3e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ||||||
Db          1 LIQLFH 6


RESULT 12
US-09-765-527-162
; Sequence 162, Application US/09765527
; Patent No. US20020006638A1
;   GENERAL INFORMATION:
;        APPLICANT: Better, Marc D.
;        TITLE OF INVENTION: Methods for Recombinant Microbial Production of
```

```
;                          Fusion Proteins and BPI-Derived Peptides
;       NUMBER OF SEQUENCES: 265
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
;            STREET: 6300 Sears Tower, 233 South Wacker Drive
;            CITY: Chicago
;            STATE: Illinois
;            COUNTRY: United States of America
;            ZIP: 60606-6402
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/765,527
;            FILING DATE: 18-Jan-2001
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 08/621,803
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Borun, Michael F.
;            REGISTRATION NUMBER: 25,447
;            REFERENCE/DOCKET NUMBER: 27129/33199
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 312/474-6300
;            TELEFAX: 312/474-0448
;            TELEX: 25-3856
;   INFORMATION FOR SEQ ID NO: 162:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 9 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: peptide
;       FEATURE:
;            NAME/KEY:  misc_feature
;            OTHER INFORMATION:  "XMP.296"
;       FEATURE:
;            NAME/KEY:  Modified-site
;            LOCATION:  C-Terminus
;            OTHER INFORMATION:  /label= Amidation
;            /note= "The C-Terminus is Amidated."
;       SEQUENCE DESCRIPTION: SEQ ID NO: 162:
US-09-765-527-162

  Query Match             100.0%;  Score 31;  DB 9;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 5.3e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ||||||
Db          4 LIQLFH 9


RESULT 13
US-09-765-527-163
; Sequence 163, Application US/09765527
; Patent No. US20020006638A1
;   GENERAL INFORMATION:
;       APPLICANT: Better, Marc D.
;       TITLE OF INVENTION: Methods for Recombinant Microbial Production of
;                           Fusion Proteins and BPI-Derived Peptides
;       NUMBER OF SEQUENCES: 265
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
;            STREET: 6300 Sears Tower, 233 South Wacker Drive
;            CITY: Chicago
;            STATE: Illinois
;            COUNTRY: United States of America
;            ZIP: 60606-6402
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/765,527
;            FILING DATE: 18-Jan-2001
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 08/621,803
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Borun, Michael F.
;            REGISTRATION NUMBER: 25,447
;            REFERENCE/DOCKET NUMBER: 27129/33199
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 312/474-6300
;            TELEFAX: 312/474-0448
;            TELEX: 25-3856
;   INFORMATION FOR SEQ ID NO: 163:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 9 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: peptide
;       FEATURE:
;            NAME/KEY:  misc_feature
;            OTHER INFORMATION:  "XMP.297"
;       FEATURE:
;            NAME/KEY:  Modified-site
;            LOCATION:  C-Terminus
;            OTHER INFORMATION:  /label= Amidation
;            /note= "The C-Terminus is Amidated."
;       SEQUENCE DESCRIPTION: SEQ ID NO: 163:
US-09-765-527-163

  Query Match             100.0%;  Score 31;  DB 9;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 5.3e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ||||||
Db          3 LIQLFH 8


RESULT 14
US-09-765-527-164
; Sequence 164, Application US/09765527
; Patent No. US20020006638A1
;   GENERAL INFORMATION:
;       APPLICANT: Better, Marc D.
;       TITLE OF INVENTION: Methods for Recombinant Microbial Production of
;                           Fusion Proteins and BPI-Derived Peptides
;       NUMBER OF SEQUENCES: 265
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
;            STREET: 6300 Sears Tower, 233 South Wacker Drive
;            CITY: Chicago
;            STATE: Illinois
;            COUNTRY: United States of America
;            ZIP: 60606-6402
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/765,527
;            FILING DATE: 18-Jan-2001
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 08/621,803
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Borun, Michael F.
;            REGISTRATION NUMBER: 25,447
;            REFERENCE/DOCKET NUMBER: 27129/33199
```

```
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 312/474-6300
;            TELEFAX: 312/474-0448
;            TELEX: 25-3856
;   INFORMATION FOR SEQ ID NO: 164:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 9 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: peptide
;       FEATURE:
;            NAME/KEY:  misc_feature
;            OTHER INFORMATION:  "XMP.298"
;       FEATURE:
;            NAME/KEY:  Modified-site
;            LOCATION:  C-Terminus
;            OTHER INFORMATION:  /label= Amidation
;            /note= "The C-Terminus is Amidated."
;       SEQUENCE DESCRIPTION: SEQ ID NO: 164:
US-09-765-527-164

  Query Match             100.0%;  Score 31;  DB 9;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 5.3e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ||||||
Db          2 LIQLFH 7


RESULT 15
US-09-881-490-129
; Sequence 129, Application US/09881490
; Patent No. US20020077298A1
;   GENERAL INFORMATION:
;       APPLICANT: Little II, Roger G.
;                  Lim, Edward
;                  Fadem, Mitchell B.
;       TITLE OF INVENTION: Anti-Fungal Peptides
;       NUMBER OF SEQUENCES: 211
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: McAndrews, Held & Malloy, Ltd.
;            STREET: 500 West Madison Street, 34th FloorDrive
;            CITY: Chicago
;            STATE: Illinois
;            COUNTRY: United States of America
;            ZIP: 60661
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/881,490
;            FILING DATE: 14-Jun-2001
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/119,858
;            FILING DATE: <Unknown>
;            APPLICATION NUMBER: 08/372,105
;            FILING DATE: 13-JAN-95
;            APPLICATION NUMBER: 08/306,473
;            FILING DATE: 15-SEP-94
;            APPLICATION NUMBER: 08/273,540
;            FILING DATE: 11-JUL-94
;            APPLICATION NUMBER: 08/209,762
;            FILING DATE: 11-MAR-94
;            APPLICATION NUMBER: 08/183,222
;            FILING DATE: 14-JAN-94
;            APPLICATION NUMBER: 08/093,202
;            FILING DATE: 15-JUL-93
;            APPLICATION NUMBER: 08/030,644
;            FILING DATE: 12-MAR-93
;       ATTORNEY/AGENT INFORMATION:
;            NAME: McNicholas, Janet M.
;            REGISTRATION NUMBER: 32,918
;            REFERENCE/DOCKET NUMBER: 100-238/11021US01
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 312/707-8889
;            TELEFAX: 312/707-9155
;            TELEX: 650 388-1248
;   INFORMATION FOR SEQ ID NO: 129:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 9 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: peptide
;       FEATURE:
;            NAME/KEY:  misc_feature
;            OTHER INFORMATION:  "XMP.296"
;       FEATURE:
;            NAME/KEY:  Modified-site
;            LOCATION:  C-Terminus
;            OTHER INFORMATION:  /label= Amidation
;            /note= "The C-Terminus is Amidated"
;       SEQUENCE DESCRIPTION: SEQ ID NO: 129:
US-09-881-490-129

  Query Match             100.0%;  Score 31;  DB 9;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 5.3e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ||||||
Db          4 LIQLFH 9


Search completed: October 14, 2003, 14:44:49
Job time : 15.625 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 14:33:18 ; Search time 7.875 Seconds
(without alignments)
73.271 Million cell updates/sec

Title: US-09-881-490-150
Perfect score: 31
Sequence: 1 LIQLFH 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	31	100.0	146	2	B86721	transcription regu
2	31	100.0	250	4	S43383	bactericidal/perme
3	31	100.0	487	2	A30909	bactericidal/perme
4	31	100.0	898	2	A98048	H+/K+-exchanging A
5	31	100.0	914	2	E95180	cation-transportin
6	30	96.8	1039	2	T30856	protein F2 - Strep
7	29	93.5	189	1	LPHUD	apolipoprotein D p
8	29	93.5	313	2	S38538	aspartoacylase - h
9	29	93.5	370	2	D70076	ethanolamine trans
10	28	90.3	179	2	D83660	hypoxanthine-guani
11	28	90.3	212	2	S71432	farnesyl-pyrophosp
12	28	90.3	339	2	T42081	farnesyl-pyrophosp
13	28	90.3	347	2	T11664	geranyltranstransf
14	28	90.3	469	2	T51882	hypothetical prote
15	28	90.3	509	2	T32201	hypothetical prote
16	28	90.3	1277	2	S54451	hypothetical prote
17	28	90.3	2024	2	A54103	centrosome autoant
18	28	90.3	2301	2	T02323	nodulin-like prote
19	27	87.1	24	2	T42378	phosphoprotein pho
20	27	87.1	153	1	MYRB	myoglobin - rabbit
21	27	87.1	159	2	E90327	conserved hypothet
22	27	87.1	213	2	D82849	cell division prot
23	27	87.1	324	2	T27302	hypothetical prote
24	27	87.1	343	2	T32173	hypothetical prote
25	27	87.1	357	2	T22879	hypothetical prote
26	27	87.1	357	2	T22880	hypothetical prote
27	27	87.1	359	2	T27200	hypothetical prote
28	27	87.1	360	2	T24641	hypothetical prote
29	27	87.1	478	2	S25821	transposase - Baci
30	27	87.1	478	2	C59095	hypothetical prote
31	27	87.1	499	1	A37924	heparin cofactor I
32	27	87.1	648	2	T47988	serine/threonine-p
33	27	87.1	669	1	S62919	probable membrane
34	27	87.1	671	2	AE0486	probable membrane
35	27	87.1	716	2	H84421	probable receptor-
36	27	87.1	812	2	B97401	uridylyltransferas
37	27	87.1	869	2	S68428	probable chloride
38	27	87.1	928	2	T24868	hypothetical prote
39	27	87.1	942	2	AB2619	hypothetical prote
40	27	87.1	1186	2	AG1928	two-component hybr
41	27	87.1	1822	2	S44849	K12H4.8 protein -
42	27	87.1	4196	2	T43274	dynein heavy chain
43	26	83.9	33	2	C86007	hypothetical prote
44	26	83.9	70	2	C90886	hypothetical prote
45	26	83.9	70	2	D85732	hypothetical prote

ALIGNMENTS

RESULT 1
B86721
transcription regulator [imported] - Lactococcus lactis subsp. lactis (strain IL1403
C;Species: Lactococcus lactis subsp. lactis
C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C;Accession: B86721
R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; E
Genome Res. 11, 731-753, 2001
A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lacti
A;Reference number: A86625; MUID:21235186; PMID:11337471
A;Accession: B86721
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-146 <STO>
A;Cross-references: GB:AE005176; PID:g12723688; PIDN:AAK04868.1; GSPDB:GN00146
A;Experimental source: strain IL1403
C;Genetics:
A;Gene: rmaG

Query Match 100.0%; Score 31; DB 2; Length 146;
Best Local Similarity 100.0%; Pred. No. 5.6;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 LIQLFH 6
              ||||||
Db         12 LIQLFH 17
```

RESULT 2
S43383
bactericidal/permeability-increasing protein - synthetic
C;Species: synthetic
A;Note: Homo sapiens (man) gene engineered and expressed in Escherichia coli
C;Date: 20-Oct-1994 #sequence_revision 15-Feb-1996 #text_change 15-Feb-1996
C;Accession: S43383
R;Qi, S.Y.; Li, Y.; O'Connor, C.D.
Biochem. J. 298, 711-718, 1994
A;Title: The region around residue 115 of human bactericidal/permeability-increasing
of a gene coding for the active domain and characterization of recombinant proteins.
A;Reference number: S43383
A;Accession: S43383
A;Molecule type: DNA
A;Residues: 1-250 <QIS>

Query Match 100.0%; Score 31; DB 4; Length 250;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 LIQLFH 6
              ||||||
Db        155 LIQLFH 160
```

RESULT 3
A30909
bactericidal/permeability-increasing protein precursor - human
N;Alternate names: 55K bactericidal protein
C;Species: Homo sapiens (man)
C;Date: 18-Apr-1989 #sequence_revision 18-Apr-1989 #text_change 20-Aug-1999
C;Accession: A33850; B54136; A29464; A43600; A49716; A30909
R;Gray, P.W.; Flaggs, G.; Leong, S.R.; Gumina, R.J.; Weiss, J.; Ooi, C.E.; Elsbach, P.
J. Biol. Chem. 264, 9505-9509, 1989
A;Title: Cloning of the cDNA of a human neutrophil bactericidal protein. Structural and
A;Reference number: A33850; MUID:89255455; PMID:2722846
A;Accession: A33850
A;Molecule type: mRNA
A;Residues: 1-487 <GRA>
A;Cross-references: GB:J04739; NID:g179528; PIDN:AAA51841.1; PID:g179529
R;Wilde, C.G.; Seilhamer, J.J.; McGrogan, M.; Ashton, N.; Snable, J.L.; Lane, J.C.; Leon
J. Biol. Chem. 269, 17411-17416, 1994
A;Title: Bactericidal/permeability-increasing protein and lipopolysaccharide (LPS)-bindi
A;Reference number: A54136; MUID:94292492; PMID:7517398
A;Accession: B54136
A;Status: nucleic acid sequence not shown; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-374,'L',376-487 <WIL>
A;Experimental source: HL-60 cells
A;Note: sequence extracted from NCBI backbone (NCBIP:149855)
R;Ooi, C.E.; Weiss, J.; Elsbach, P.; Frangione, B.; Mannion, B.
J. Biol. Chem. 262, 14891-14894, 1987
A;Title: A 25-kDa amino-terminal fragment carries all the antibacterial activities of th
A;Reference number: A29464; MUID:88033057; PMID:3667613
A;Accession: A29464
A;Molecule type: protein
A;Residues: 32-51 <OOI>
A;Experimental source: neutrophils
R;Wasiluk, K.R.; Skubitz, K.M.; Gray, B.H.
Infect. Immun. 59, 4193-4200, 1991
A;Title: Comparison of granule proteins from human polymorphonuclear leukocytes which ar
A;Reference number: A43600; MUID:92040097; PMID:1937776
A;Accession: A43600
A;Molecule type: protein
A;Residues: 32-52,'R' <WAS>
R;Little, R.G.; Kelner, D.N.; Lim, E.; Burke, D.J.; Conlon, P.J.
J. Biol. Chem. 269, 1865-1872, 1994
A;Title: Functional domains of recombinant bactericidal/permeability increasing protein
A;Reference number: A49716; MUID:94124531; PMID:8294435
A;Accession: A49716
A;Molecule type: protein
A;Residues: 32-130;132-141;143-165;202-215,'E',217-225 <LIT>
C;Comment: The bactericidal/permeability-increasing protein (BPI) is a 60 kD membrane-as
which is specific for gram-negative bacteria. BPI has a high affinity for lipopolysaccha
between BPI and an LPS-binding protein from liver and cholesteryl ester transfer protein
C;Genetics:
A;Gene: GDB:BPI
A;Cross-references: GDB:131572; OMIM:109195
A;Map position: 20q11.23-20q12
C;Superfamily: lipopolysaccharide-binding protein
C;Keywords: antibacterial; cytotoxin; glycoprotein; heparin binding; neutrophil
F;1-31/Domain: signal sequence #status predicted <SIG>
F;32-487/Product: bactericidal permeability-increasing protein #status predicted <MAT>
F;32-51/Region: bactericidal #status predicted
F;380/Binding site: carbohydrate (Asn) (covalent) #status predicted

 Query Match 100.0%; Score 31; DB 2; Length 487;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LIQLFH 6
||||||
Db 185 LIQLFH 190

RESULT 4
A98048
H+/K+-exchanging ATPase (EC 3.6.3.10) [imported] - Streptococcus pneumoniae (strain
C;Species: Streptococcus pneumoniae
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 19-Apr-2002
C;Accession: A98048
R;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszczak, L.; Burgett, S.; DeHoff, B.S
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S
A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A;Reference number: A97872; MUID:21429245; PMID:11544234
A;Accession: A98048
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-898 <KUR>
A;Cross-references: GB:AE007317; PIDN:AAL00214.1; PID:g15459062; GSPDB:GN00174
C;Genetics:
A;Gene: pacL
C;Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding dom
C;Keywords: hydrolase

 Query Match 100.0%; Score 31; DB 2; Length 898;
 Best Local Similarity 100.0%; Pred. No. 39;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LIQLFH 6
||||||
Db 806 LIQLFH 811

RESULT 5
E95180
cation-transporting ATPase, E1-E2 family SP1551 [imported] - Streptococcus pneumonia
C;Species: Streptococcus pneumoniae
C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 24-Aug-2001
C;Accession: E95180
R;Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; I
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapp
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morris
A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A;Reference number: A95000; MUID:21357209; PMID:11463916
A;Accession: E95180
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-914 <KUR>
A;Cross-references: GB:AE005672; PIDN:AAK75638.1; PID:g14973040; GSPDB:GN00164; TIGR
A;Experimental source: strain TIGR4
C;Genetics:
A;Gene: SP1551
C;Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding dom

 Query Match 100.0%; Score 31; DB 2; Length 914;
 Best Local Similarity 100.0%; Pred. No. 40;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LIQLFH 6
||||||
Db 822 LIQLFH 827

RESULT 6
T30856
protein F2 - Streptococcus pyogenes
C;Species: Streptococcus pyogenes
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 11-May-2000
C;Accession: T30856
R;Jaffe, J.; Natanson-Yaron, S.; Caparon, M.G.; Hanski, E.
Mol. Microbiol. 21, 373-384, 1996
A;Title: Protein F2, a novel fibronectin-binding protein from Streptococcus pyogenes,
A;Reference number: Z20907; MUID:97011581; PMID:8858591
A;Accession: T30856

```
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1039 <JAP>
A;Cross-references: EMBL:U31980; NID:g1654115; PID:g1654116; PIDN:AAC44522.1
C;Genetics:
A;Note: PRTF2

  Query Match             96.8%;  Score 30;  DB 2;  Length 1039;
  Best Local Similarity   83.3%;  Pred. No. 79;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              |:||||
Db        972 LVQLFH 977


RESULT 7
LPHUD
apolipoprotein D precursor [validated] - human
N;Alternate names: apoD; gross cystic disease fluid protein 24 (GCDFP-24)
C;Species: Homo sapiens (man)
C;Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 15-Sep-2000
C;Accession: A26958; A03224; S13050
R;Drayna, D.T.; McLean, J.W.; Wion, K.L.; Trent, J.M.; Drabkin, H.A.; Lawn, R.M.
DNA 6, 199-204, 1987
A;Title: Human apolipoprotein D gene: gene sequence, chromosome localization, and homolo
A;Reference number: A26958; MUID:87246069; PMID:2439269
A;Accession: A26958
A;Molecule type: DNA
A;Residues: 1-189 <DRA1>
A;Cross-references: GB:M16648; GB:M16649; GB:M16695; GB:M16696; NID:g178845; PIDN:AAA517
R;Drayna, D.; Fielding, C.; McLean, J.; Baer, B.; Castro, G.; Chen, E.; Comstock, L.; He
J. Biol. Chem. 261, 16535-16539, 1986
A;Title: Cloning and expression of human apolipoprotein D cDNA.
A;Reference number: A03224; MUID:87057347; PMID:3453108
A;Accession: A03224
A;Molecule type: mRNA
A;Residues: 1-189 <DRA2>
A;Cross-references: GB:J02611; NID:g178840; PIDN:AAB59517.1; PID:g178841
A;Note: part of the sequence was confirmed by protein sequencing
R;Balbin, M.; Freije, J.M.P.; Fueyo, A.; Sanchez, L.M.; Lopez-Otin, C.
Biochem. J. 271, 803-807, 1990
A;Title: Apolipoprotein D is the major protein component in cyst fluid from women with h
A;Reference number: S13050; MUID:91058519; PMID:2244881
A;Accession: S13050
A;Molecule type: protein
A;Residues: 28-41;42-45;'X',47-51;52-60;61-64,'X',66-73;76-82;83-97,'X',99-104;109-110,'
A;Experimental source: mammary cyst fluid
A;Note: carbohydrate binding sites determined; the sequence of the amino-terminal pyrogl
R;Peitsch, M.C.; Boguski, M.S.
submitted to the Brookhaven Protein Data Bank, April 1992
A;Reference number: A50026; PDB:2APD
A;Contents: annotation; theoretical model, residues 21-189
R;Yang, C.Y.; Gu, Z.W.; Blanco-Vaca, F.; Gaskell, S.J.; Yang, M.; Massey, J.B.; Gotto Jr
Biochemistry 33, 12451-12455, 1994
A;Title: Structure of human apolipoprotein D: locations of the intermolecular and intram
A;Reference number: A55901; MUID:95001965; PMID:7918467
A;Contents: annotation; confirmation of peptide sequence overlap; disulfide bonds; N-gly
R;Schindler, P.A.; Settineri, C.A.; Collet, X.; Fielding, C.J.; Burlingame, A.L.
Protein Sci. 4, 791-803, 1995
A;Title: Site-specific detection and structural characterization of the glycosylation of
pectrometry and sequential glycosidase digestion.
A;Reference number: A57914; MUID:95338133; PMID:7613477
A;Contents: annotation; N-glycosylation
C;Comment: ApoD occurs in the macromolecular complex with lecithin-cholesterol acyltrans
C;Comment: ApoD is primarily localized in HDL (60-65%), with most of the remainder in VH
C;Comment: ApoD has been found in liver, intestine, pancreas, kidney, placenta, adrenal,
C;Genetics:
A;Gene: GDB:APOD
A;Cross-references: GDB:119690; OMIM:107740
A;Map position: 3q27-3qter
A;Introns: 41/3; 82/2; 112/1
A;Note: the first intron occurs before the initiator codon
C;Superfamily: lipocalin; lipocalin homology
C;Keywords: glycoprotein; lipid binding; lipid transport; lipoprotein; plasma; pyroq
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-189/Product: apolipoprotein D #status experimental <MAT>
F;37-185/Domain: lipocalin homology <LIP>
F;21/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status exper
F;28-134,61-185/Disulfide bonds: #status experimental
F;65,98/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;136/Disulfide bonds: interchain (to apolipoprotein A-II 29) #status experimental

  Query Match             93.5%;  Score 29;  DB 1;  Length 189;
  Best Local Similarity   83.3%;  Pred. No. 22;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              :|||||
Db        137 IIQLFH 142


RESULT 8
S38538
aspartoacylase - human
C;Species: Homo sapiens (man)
C;Date: 07-Jun-1994 #sequence_revision 17-Nov-1995 #text_change 28-May-1999
C;Accession: S38538
R;Kaul, R.; Gao, G.P.; Balamurugan, K.; Matalon, R.
Nature Genet. 5, 118-123, 1993
A;Title: Cloning of the human aspartoacylase cDNA and a common missense mutation in
A;Reference number: S38538; MUID:94073185; PMID:8252036
A;Accession: S38538
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-313 <KAU>
A;Cross-references: GB:S67156; NID:g455833; PIDN:AAB29190.1; PID:g455834

  Query Match             93.5%;  Score 29;  DB 2;  Length 313;
  Best Local Similarity   83.3%;  Pred. No. 37;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              |||:||
Db        136 LIQMFH 141


RESULT 9
D70076
ethanolamine transporter homolog yxeR - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999
C;Accession: D70076
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Be
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Ga
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo,
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardir
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Ma
y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portet
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scar
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; S
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiy
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshic
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtil
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: D70076
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-370 <KUN>
A;Cross-references: GB:Z99124; GB:AL009126; NID:g2636442; PIDN:CAB15981.1; PID:e1184
A;Experimental source: strain 168
C;Genetics:
```

A;Gene: yxeR

Query Match 93.5%; Score 29; DB 2; Length 370;
Best Local Similarity 83.3%; Pred. No. 45;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LIQLFH 6
:|||||
Db 193 MIQLFH 198

RESULT 10
D83660
hypoxanthine-guanine phosphoribosyltransferase hprT [imported] - Bacillus halodurans (st
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C;Accession: D83660
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: D83660
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-179 <STO>
A;Cross-references: GB:AP001507; GB:BA000004; NID:g10172612; PIDN:BAB03803.1; GSPDB:GN00
A;Experimental source: strain C-125
C;Genetics:
A;Gene: hprT
C;Superfamily: hypoxanthine phosphoribosyltransferase

Query Match 90.3%; Score 28; DB 2; Length 179;
Best Local Similarity 83.3%; Pred. No. 35;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LIQLFH 6
||:|||
Db 110 LIELFH 115

RESULT 11
S71432
farnesyl-pyrophosphate synthetase - Sphaceloma manihoticola (fragment)
N;Alternate names: farnesyl-diphosphate synthase; farnesylpyrophosphate synthase
N;Contains: dimethylallyltranstransferase (EC 2.5.1.1); geranyltranstransferase (EC 2.5.
C;Species: Sphaceloma manihoticola
C;Date: 12-Mar-1998 #sequence_revision 12-Mar-1998 #text_change 20-Jun-2000
C;Accession: S71432
R;Homann, V.; Mende, K.; Arntz, C.; Ilardi, V.; Macino, G.; Morelli, G.; Boese, G.; Tudz
Curr. Genet. 30, 232-239, 1996
A;Title: The isoprenoid pathway: cloning and characterization of fungal FPPS genes.
A;Reference number: S71431; MUID:96337905; PMID:8753652
A;Accession: S71432
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-212 <HOM>
A;Cross-references: EMBL:X96941; NID:g1524344; PIDN:CAA65642.1; PID:g1524345
A;Experimental source: strain Lul116
C;Genetics:
A;Gene: FPPS
C;Function:
A;Pathway: isoprenoid biosynthesis
C;Superfamily: farnesyl-pyrophosphate synthetase
C;Keywords: cytosol; dimer; isoprenoid biosynthesis; multifunctional enzyme; transferase
F;48-52/Region: DDXXD motif
F;188-192/Region: DDXXD motif

Query Match 90.3%; Score 28; DB 2; Length 212;
Best Local Similarity 83.3%; Pred. No. 42;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LIQLFH 6
||:|||
Db 101 LIELFH 106

RESULT 12
T42081
farnesyl-pyrophosphate synthetase - fission yeast (Schizosaccharomyces pombe) (fragm
N;Contains: dimethylallyltranstransferase (EC 2.5.1.1); geranyltranstransferase (EC
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jul-2000
C;Accession: T42081
R;Yoshioka, S.; Kato, K.; Nakai, K.; Okayama, H.; Nojima, H.
DNA Res. 4, 363-369, 1997
A;Title: Identification of open reading frames in Schizosaccharomyces pombe cDNAs.
A;Reference number: Z17323; MUID:98162722; PMID:9501991
A;Accession: T42081
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-339 <YOS>
A;Cross-references: EMBL:D89104; NID:g1749415; PIDN:BAA13767.1; PID:g1749416
A;Experimental source: strain PR745
C;Function:
A;Description: catalyzes the sequential condensation of isopentenyl diphosphate with
diphosphate with geranyl diphosphate forming farnesyl d
A;Pathway: isoprenoid biosynthesis
A;Note: first enzyme of isoprenoid biosynthesis
C;Superfamily: farnesyl-pyrophosphate synthetase
C;Keywords: isoprenoid biosynthesis; multifunctional enzyme; transferase

Query Match 90.3%; Score 28; DB 2; Length 339;
Best Local Similarity 83.3%; Pred. No. 69;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LIQLFH 6
||:|||
Db 133 LIELFH 138

RESULT 13
T11664
geranyltranstransferase (EC 2.5.1.10) - fission yeast (Schizosaccharomyces pombe)
N;Alternate names: farnesyl-diphosphate synthase
C;Species: Schizosaccharomyces pombe
C;Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 20-Jun-2000
C;Accession: T11664
R;Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, August 1997
A;Reference number: Z17305
A;Accession: T11664
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-347 <BAR>
A;Cross-references: EMBL:Z98533
C;Superfamily: farnesyl-pyrophosphate synthetase
C;Keywords: transferase

Query Match 90.3%; Score 28; DB 2; Length 347;
Best Local Similarity 83.3%; Pred. No. 71;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LIQLFH 6
||:|||
Db 148 LIELFH 153

RESULT 14
T51882
hypothetical protein DKFZp547I194.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 03-Nov-2000
C;Accession: T51882
R;Bloecker, H.; Boecher, M.; Brandt, P.; Mewes, H.W.; Weil, B.; Wiemann, S.
submitted to the Protein Sequence Database, July 2000
A;Reference number: Z25853

A;Accession: T51882
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-469 <AAA>
A;Cross-references: EMBL:AL390171
A;Experimental source: fetal brain; clone DKFZp5471194
C;Genetics:
A;Note: DKFZp5471194.1
C;Superfamily: Caenorhabditis elegans hypothetical protein F08C6.3

Query Match 90.3%; Score 28; DB 2; Length 469;
Best Local Similarity 83.3%; Pred. No. 98;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LIQLFH 6
||||:|
Db 427 LIQLYH 432

RESULT 15
T32201
hypothetical protein T02B11.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C;Accession: T32201
R;Goela, D.
submitted to the EMBL Data Library, September 1997
A;Description: The sequence of C. elegans cosmid T02B11.
A;Reference number: Z21135
A;Accession: T32201
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-509 <GOE>
A;Cross-references: EMBL:AF022979; PIDN:AAB69908.1; GSPDB:GN00023; CESP:T02B11.2
A;Experimental source: strain Bristol N2; clone T02B11
C;Genetics:
A;Gene: CESP:T02B11.2
A;Map position: 5
A;Introns: 149/2; 171/3; 302/1; 316/1; 387/3; 497/1

Query Match 90.3%; Score 28; DB 2; Length 509;
Best Local Similarity 83.3%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LIQLFH 6
||:|||
Db 313 LIELFH 318

Search completed: October 14, 2003, 14:37:59
Job time : 10.875 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 14:33:18 ; Search time 4.25 Seconds
(without alignments)
66.391 Million cell updates/sec

Title: US-09-881-490-150
Perfect score: 31
Sequence: 1 LIQLFH 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	31	100.0	483	1	BPI_HUMAN	P17213 homo sapien
2	29	93.5	189	1	APOD_HUMAN	P05090 homo sapien
3	29	93.5	313	1	ACY2_BOVIN	P46446 bos taurus
4	29	93.5	313	1	ACY2_HUMAN	P45381 homo sapien
5	29	93.5	370	1	YXER_BACSU	P54957 bacillus su
6	28	90.3	347	1	FPPS_SCHPO	O14230 schizosacch
7	28	90.3	1277	1	PDS5_YEAST	Q04264 saccharomyc
8	27	87.1	153	1	MYG_RABIT	P02170 oryctolagus
9	27	87.1	359	1	RT22_MOUSE	Q9cxw2 mus musculu
10	27	87.1	360	1	RT22_HUMAN	P82650 homo sapien
11	27	87.1	416	1	PROA_LEPIN	P94872 leptospira
12	27	87.1	478	1	T23E_BACTF	Q02403 bacillus th
13	27	87.1	499	1	HEP2_HUMAN	P05546 homo sapien
14	27	87.1	669	1	YNA8_YEAST	P53983 saccharomyc
15	27	87.1	869	1	CLC6_HUMAN	P51797 homo sapien
16	27	87.1	870	1	CLC6_MOUSE	O35454 mus musculu
17	27	87.1	942	1	GLND_AGRT5	Q8uif1 agrobacteri
18	27	87.1	1845	1	DCR1_CAEEL	P34529 caenorhabdi
19	27	87.1	4196	1	DYHC_SCHPO	O13290 schizosacch
20	26	83.9	101	1	S107_BOVIN	Q28050 bos taurus
21	26	83.9	151	1	RS15_PICMA	O65059 picea maria
22	26	83.9	153	1	RS15_ELAOL	Q945u1 elaeis olei
23	26	83.9	168	1	YPQE_BACSU	P50829 bacillus su
24	26	83.9	206	1	YI87_PYRHO	O59533 pyrococcus
25	26	83.9	289	1	YF03_PASMU	Q9ckv2 pasteurella
26	26	83.9	304	1	LDCA_ECO57	Q8xdj8 escherichia
27	26	83.9	304	1	LDCA_ECOL6	P59238 escherichia
28	26	83.9	304	1	LDCA_ECOLI	P76008 escherichia
29	26	83.9	309	1	FN3X_HUMAN	Q9ha64 homo sapien
30	26	83.9	320	1	YC55_PORPU	P51203 porphyra pu
31	26	83.9	349	1	FPPS_KLULA	P49349 kluyveromyc
32	26	83.9	372	1	VGLI_HSV23	P06764 herpes simp
33	26	83.9	372	1	VGLI_HSV2H	P13291 herpes simp
34	26	83.9	506	1	PCCB_BACSU	P54541 bacillus su
35	26	83.9	507	1	CDC7_YEAST	P06243 saccharomyc
36	26	83.9	540	1	THI6_YEAST	P41835 s thiamine
37	26	83.9	611	1	YJS1_YEAST	P46987 saccharomyc
38	26	83.9	642	1	UL77_HCMVA	P16726 human cytom
39	26	83.9	678	1	PTR3_YEAST	P43606 saccharomyc
40	26	83.9	1257	1	CCAA_BACTU	Q45754 bacillus th
41	26	83.9	1451	1	A2M2_MOUSE	P28666 mus musculu
42	26	83.9	1476	1	A2M1_MOUSE	P28665 mus musculu
43	26	83.9	2778	1	FAF_DROME	P55824 drosophila
44	26	83.9	3655	1	YAMB_SCHPO	Q10064 schizosacch
45	25	80.6	67	1	CD52_CANFA	Q28896 canis famil

ALIGNMENTS

```
RESULT 1
BPI_HUMAN
ID   BPI_HUMAN      STANDARD;      PRT;   483 AA.
AC   P17213; Q9BYZ9; Q9H1L2; Q9H1M8; Q9H203; Q9UD65;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Bactericidal permeability-increasing protein precursor (BPI) (CAP 57).
GN   BPI.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 28-64.
RX   MEDLINE=89255455; PubMed=2722846;
RA   Gray P.W., Flaggs G., Leong S.R., Gumina R.J., Weiss J., Ooi C.E.,
RA   Elsbach P.;
RT   "Cloning of the cDNA of a human neutrophil bactericidal protein.
RT   Structural and functional correlations.";
RL   J. Biol. Chem. 264:9505-9509(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94292492; PubMed=7517398;
RA   Wilde C.G., Seilhamer J.J., McGrogan M., Ashton N., Snable J.L.,
RA   Lane J.C., Leong S.R., Thornton M.B., Miller K.L., Scott R.W.;
RT   "Bactericidal/permeability-increasing protein and lipopolysaccharide
RT   (LPS)-binding protein. LPS binding properties and effects on LPS-
RT   mediated cell activation.";
RL   J. Biol. Chem. 269:17411-17416(1994).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Xu J., Wang H.;
RT   "Cloning of cDNA of human bactericidal/permeability-increasing
RT   protein.";
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21638749; PubMed=11780052;
RA   Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA   Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA   Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA   Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA   Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA   Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA   Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA   Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
RA   Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA   Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA   Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA   Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA   Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA   Lehvaeslaiho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA   Marsh V.L., Martin S.L., McConnachie L.J., McLay K., McMurray A.A.,
RA   Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA   Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
```

```
RA   Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA   Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,
RA   Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA   Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA   Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA   Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
RA   Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA   Rogers J.;
RT   "The DNA sequence and comparative analysis of human chromosome 20.";
RL   Nature 414:865-871(2001).
RN   [5]
RP   SEQUENCE OF 28-42.
RX   MEDLINE=88033057; PubMed=3667613;
RA   Ooi C.E., Weiss J., Elsbach P., Frangione B., Mannion B.;
RT   "A 25-kDa NH2-terminal fragment carries all the antibacterial
RT   activities of the human neutrophil 60-kDa
RT   bactericidal/permeability-increasing protein.";
RL   J. Biol. Chem. 262:14891-14894(1987).
RN   [6]
RP   SEQUENCE OF 28-47.
RX   MEDLINE=89315847; PubMed=2501794;
RA   Gabay J.E., Scott R.W., Campanelli D., Griffith J., Wilde C.,
RA   Marra M.N., Seeger M., Nathan C.F.;
RT   "Antibiotic proteins of human polymorphonuclear leukocytes.";
RL   Proc. Natl. Acad. Sci. U.S.A. 86:5610-5614(1989).
RN   [7]
RP   X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
RX   MEDLINE=97334442; PubMed=9186532;
RA   Beamer L.J., Carroll S.F., Eisenberg D.;
RT   "Crystal structure of human BPI and two bound phospholipids at 2.4-A
RT   resolution.";
RL   Science 276:1861-1864(1997).
CC   -!- FUNCTION: THE CYTOTOXIC ACTION OF BPI IS LIMITED TO MANY SPECIES
CC       OF GRAM-NEGATIVE BACTERIA; THIS SPECIFICITY MAY BE EXPLAINED BY A
CC       STRONG AFFINITY OF THE VERY BASIC N-TERMINAL HALF FOR THE
CC       NEGATIVELY CHARGED LIPOPOLYSACCHARIDES THAT ARE UNIQUE TO THE
CC       GRAM-NEGATIVE BACTERIAL OUTER ENVELOPE.
CC   -!- SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED IN POLYMORPHONUCLEAR
CC       LEUKOCYTES (PMN) GRANULES.
CC   -!- TISSUE SPECIFICITY: RESTRICTED TO CELLS OF THE MYELOID SERIES.
CC   -!- DOMAIN: THE N-TERMINAL REGION MAY BE EXPOSED TO THE INTERIOR OF
CC       THE GRANULE, WHEREAS THE C-TERMINAL PORTION MAY BE EMBEDDED IN THE
CC       MEMBRANE. DURING PHAGOCYTOSIS AND DEGRANULATION, PROTEASES MAY BE
CC       RELEASED AND ACTIVATED AND CLEAVE BPI AT THE JUNCTION OF THE N-
CC       AND C-TERMINAL PORTIONS OF THE MOLECULE, PROVIDING CONTROLLED
CC       RELEASE OF THE N-TERMINAL ANTIBACTERIAL FRAGMENT WHEN BACTERIA ARE
CC       INGESTED.
CC   -!- SIMILARITY: BELONGS TO THE BPI/CETP/LBP/PLTP FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J04739; AAA51841.1; ALT_INIT.
DR   EMBL; AF322588; AAG42844.1; -.
DR   EMBL; AL359555; CAC13043.1; -.
DR   EMBL; AL499625; CAC27350.1; -.
DR   EMBL; AL391692; CAC10453.1; -.
DR   PDB; 1BP1; 04-SEP-97.
DR   PDB; 1EWF; 21-JUN-00.
DR   Genew; HGNC:1095; BPI.
DR   MIM; 109195; -.
DR   GO; GO:0005887; C:integral to plasma membrane; TAS.
DR   InterPro; IPR001124; LBP_BPI_CETP.
DR   Pfam; PF01273; LBP_BPI_CETP; 1.
DR   Pfam; PF02886; LBP_BPI_CETP_C; 1.
DR   SMART; SM00328; BPI1; 1.
DR   SMART; SM00329; BPI2; 1.
DR   PROSITE; PS00400; LBP_BPI_CETP; 1.
KW   Antibiotic; Signal; Transmembrane; Glycoprotein; 3D-structure.
FT   SIGNAL       1    27
FT   CHAIN       28   483       BACTERICIDAL PERMEABILITY-INCREASING
FT                               PROTEIN.
FT   SITE        236   241       CLEAVAGE SITES FOR ELASTASE (POTENTIAL).
FT   TRANSMEM    365   385       POTENTIAL.
FT   CONFLICT     12    12       V -> A (IN REF. 3 AND 4).
FT   CONFLICT    212   212       K -> E (IN REF. 4).
FT   CONFLICT    351   351       P -> S (IN REF. 3).
FT   CONFLICT    371   371       F -> L (IN REF. 2).
FT   CONFLICT    400   400       N -> D (IN REF. 3).
FT   CONFLICT    407   407       K -> R (IN REF. 3).
FT   STRAND       32    37
FT   HELIX        38    56
FT   TURN         57    58
FT   STRAND       64    70
FT   TURN         71    73
FT   STRAND       74    89
FT   STRAND       93    98
FT   TURN         99   101
FT   STRAND      102   122
FT   TURN        123   124
FT   STRAND      125   149
FT   TURN        150   153
FT   STRAND      154   163
FT   STRAND      168   172
FT   TURN        174   175
FT   HELIX       179   188
FT   TURN        189   189
FT   HELIX       190   217
FT   TURN        218   218
FT   STRAND      223   225
FT   STRAND      230   233
FT   STRAND      236   236
FT   STRAND      241   242
FT   STRAND      246   251
FT   STRAND      254   257
FT   STRAND      281   286
FT   HELIX       287   299
FT   TURN        300   301
FT   STRAND      304   307
FT   HELIX       309   311
FT   STRAND      321   321
FT   HELIX       322   325
FT   TURN        326   328
FT   TURN        330   331
FT   HELIX       332   335
FT   STRAND      340   347
FT   STRAND      352   356
FT   TURN        357   358
FT   STRAND      359   363
FT   STRAND      365   373
FT   TURN        375   376
FT   STRAND      379   388
FT   STRAND      391   398
FT   TURN        399   400
FT   STRAND      401   408
FT   STRAND      412   412
FT   STRAND      415   418
FT   TURN        419   420
FT   HELIX       425   428
FT   HELIX       429   450
FT   STRAND      452   453
FT   TURN        458   459
FT   STRAND      460   470
FT   TURN        471   472
FT   STRAND      473   483
SQ   SEQUENCE   483 AA;  53396 MW;  AD58C92BCAD8F47C CRC64;

  Query Match           100.0%;  Score 31;  DB 1;  Length 483;
  Best Local Similarity  100.0%;  Pred. No. 8;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 LIQLFH 6
              ||||||
Db        181 LIQLFH 186


RESULT 2
APOD_HUMAN
ID   APOD_HUMAN     STANDARD;      PRT;   189 AA.
AC   P05090;
DT   13-AUG-1987 (Rel. 05, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Apolipoprotein D precursor (Apo-D) (ApoD).
GN   APOD.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87057347; PubMed=3453108;
RA   Drayna D.T., Fielding C., McLean J.W., Baer B., Castro G., Chen E.,
RA   Comstock L., Henzel W., Kohr W., Rhee L., Wion K.L., Lawn R.M.;
RT   "Cloning and expression of human apolipoprotein D cDNA.";
RL   J. Biol. Chem. 261:16535-16539(1986).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87246069; PubMed=2439269;
RA   Drayna D.T., McLean J.W., Wion K.L., Trent J.M., Drabkin H.A.,
RA   Lawn R.M.;
RT   "Human apolipoprotein D gene: gene sequence, chromosome localization,
RT   and homology to the alpha 2u-globulin superfamily.";
RL   DNA 6:199-204(1987).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Skin;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [4]
RP   SEQUENCE OF 21-189, DISULFIDE BONDS, AND CARBOHYDRATE-LINKAGE SITES.
RC   TISSUE=Plasma;
RX   MEDLINE=95001965; PubMed=7918467;
RA   Yang C.-Y., Gu Z.-W., Blanco-Vaca F., Gaskell S.J., Yang M.,
RA   Massey J.B., Gotto A.M. Jr., Pownall H.J.;
RT   "Structure of human apolipoprotein D: locations of the intermolecular
RT   and intramolecular disulfide links.";
RL   Biochemistry 33:12451-12455(1994).
RN   [5]
RP   SEQUENCE OF 92-189 FROM N.A.
RC   TISSUE=Lacrimal gland;
RX   MEDLINE=96148565; PubMed=8549691;
RA   Holzfeind P., Merschak P., Dieplinger H., Redl B.;
RT   "The human lacrimal gland synthesizes apolipoprotein D mRNA in
RT   addition to tear prealbumin mRNA, both species encoding members of
RT   the lipocalin superfamily.";
RL   Exp. Eye Res. 61:495-500(1995).
RN   [6]
RP   PARTIAL SEQUENCE OF 128-187.
RX   MEDLINE=91058519; PubMed=2244881;
RA   Balbin M., Freije J.M.P., Fueyo A., Sanchez L.M., Lopez-Otin C.;
RT   "Apolipoprotein D is the major protein component in cyst fluid from
RT   women with human breast gross cystic disease.";
RL   Biochem. J. 271:803-807(1990).
RN   [7]
RP   3D-STRUCTURE MODELING, AND BILIN-BINDING.
RX   MEDLINE=91190834; PubMed=2083249;
RA   Peitsch M.C., Boguski M.S.;
RT   "Is apolipoprotein D a mammalian bilin-binding protein?";
RL   New Biol. 2:197-206(1990).
RN   [8]
RP   CARBOHYDRATE-LINKAGE SITES.
RX   MEDLINE=95338133; PubMed=7613477;
RA   Schindler P.A., Settineri C.A., Collet X., Fielding C.J.,
RA   Burlingame A.L.;
RT   "Site-specific detection and structural characterization of the
RT   glycosylation of human plasma proteins lecithin:cholesterol
RT   acyltransferase and apolipoprotein D using HPLC/electrospray mass
RT   spectrometry and sequential glycosidase digestion.";
RL   Protein Sci. 4:791-803(1995).
RN   [9]
RP   VARIANTS SER-15; LEU-115 AND LYS-178.
RX   MEDLINE=99318093; PubMed=10391209;
RA   Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
RA   Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
RA   Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,
RA   Lander E.S.;
RT   "Characterization of single-nucleotide polymorphisms in coding regions
RT   of human genes.";
RL   Nat. Genet. 22:231-238(1999).
RN   [10]
RP   ERRATUM.
RA   Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
RA   Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
RA   Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,
RA   Lander E.S.;
RL   Nat. Genet. 23:373-373(1999).
CC   -!- FUNCTION: APOD OCCURS IN THE MACROMOLECULAR COMPLEX WITH LECITHIN-
CC       CHOLESTEROL ACYLTRANSFERASE. IT IS PROBABLY INVOLVED IN THE
CC       TRANSPORT AND BINDING OF BILIN. APPEARS TO BE ABLE TO TRANSPORT A
CC       VARIETY OF LIGANDS IN A NUMBER OF DIFFERENT CONTEXTS.
CC   -!- SUBUNIT: Homodimer. In plasma, also exists as a disulfide-linked
CC       heterodimer with APOA2.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed in liver, intestine, pancreas,
CC       kidney, placenta, adrenal, spleen, fetal brain tissue and tears.
CC   -!- MISCELLANEOUS: APOD IS PRIMARILY LOCALIZED IN HDL (60-65%), WITH
CC       MOST OF THE REMAINDER IN VHDL AND ONLY TRACE AMOUNTS IN VLDL AND
CC       LDL.
CC   -!- SIMILARITY: Belongs to the lipocalin family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J02611; AAB59517.1; -.
DR   EMBL; M16696; AAA51764.1; -.
DR   EMBL; BC007402; AAH07402.1; -.
DR   EMBL; S80440; AAB35919.1; -.
DR   PIR; A26958; LPHUD.
DR   PDB; 2APD; 15-OCT-94.
DR   SWISS-2DPAGE; P05090; HUMAN.
DR   Genew; HGNC:612; APOD.
DR   MIM; 107740; -.
```

```
DR   GO; GO:0005615; C:extracellular space; TAS.
DR   GO; GO:0005321; F:high-density lipoprotein; TAS.
DR   GO; GO:0006629; P:lipid metabolism; TAS.
DR   InterPro; IPR002345; Lipocalin.
DR   InterPro; IPR000566; Lipocln_cytFABP.
DR   Pfam; PF00061; lipocalin; 1.
DR   PRINTS; PR00179; LIPOCALIN.
DR   PROSITE; PS00213; LIPOCALIN; 1.
KW   Glycoprotein; Plasma; Lipid-binding; Transport; Signal; Lipocalin;
KW   3D-structure; Polymorphism; Pyrrolidone carboxylic acid.
FT   SIGNAL       1     20
FT   CHAIN       21    189       APOLIPOPROTEIN D.
FT   MOD_RES      21     21       PYRROLIDONE CARBOXYLIC ACID.
FT   DISULFID     28    134
FT   DISULFID     61    185
FT   DISULFID    136    136       INTERCHAIN (WITH HDL-ASSOCIATED
FT                                APOLIPOPROTEIN AII).
FT   CARBOHYD     65     65       N-LINKED (GLCNAC. . .).
FT   CARBOHYD     98     98       N-LINKED (GLCNAC. . .).
FT   VARIANT      15     15       F -> S (IN dbSNP:5952).
FT                                /FTId=VAR_011931.
FT   VARIANT     115    115       S -> L (IN dbSNP:5954).
FT                                /FTId=VAR_011932.
FT   VARIANT     178    178       T -> K (IN dbSNP:5955;.
FT                                /FTId=VAR_011933.
FT   STRAND       23     25
FT   HELIX        39     42
FT   STRAND       44     51
FT   HELIX        55     57
FT   STRAND       60     67
FT   TURN         70     71
FT   STRAND       75     81
FT   TURN         84     85
FT   STRAND       87     96
FT   STRAND      103    107
FT   STRAND      109    109
FT   TURN        110    111
FT   STRAND      112    112
FT   STRAND      115    123
FT   STRAND      128    137
FT   TURN        138    140
FT   STRAND      141    150
FT   HELIX       157    167
FT   HELIX       174    176
FT   STRAND      178    179
SQ   SEQUENCE   189 AA;  21275 MW;  0EAA6DE03D5E71A8 CRC64;

  Query Match             93.5%;  Score 29;  DB 1;  Length 189;
  Best Local Similarity   83.3%;  Pred. No. 8.6;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              :|||||
Db        137 IIQLFH 142


RESULT 3
ACY2_BOVIN
ID   ACY2_BOVIN     STANDARD;      PRT;   313 AA.
AC   P46446;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   Aspartoacylase (EC 3.5.1.15) (Aminoacylase-2) (ACY-2).
GN   ASPA OR ACY2.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX   MEDLINE=94073185; PubMed=8252036;
RA   Kaul R., Gao G.P., Balamurugan K., Matalon R.;
RT   "Cloning of the human aspartoacylase cDNA and a common missense
RT   mutation in Canavan disease.";
RL   Nat. Genet. 5:118-123(1993).
CC   -!- FUNCTION: RESPONSIBLE FOR THE CLEAVAGE OF N-ACETYLASPARTIC ACID
CC       WHICH OCCURS IN HIGH CONCENTRATION IN BRAIN.
CC   -!- CATALYTIC ACTIVITY: N-acyl-L-aspartate + H(2)O = a fatty acid
CC       anion + L-aspartate.
CC   -!- SIMILARITY: TO HUMAN ASPARTOACYLASE.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; S74726; AAD14980.1; -.
DR   InterPro; IPR007036; Aste_AspA.
DR   Pfam; PF04952; Aste_AspA; 1.
KW   Hydrolase.
FT   ACT_SITE     21     21       POTENTIAL.
FT   ACT_SITE    285    285       POTENTIAL.
SQ   SEQUENCE   313 AA;  35738 MW;  1414D37921891009 CRC64;

  Query Match             93.5%;  Score 29;  DB 1;  Length 313;
  Best Local Similarity   83.3%;  Pred. No. 15;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              |||:||
Db        136 LIQMFH 141


RESULT 4
ACY2_HUMAN
ID   ACY2_HUMAN     STANDARD;      PRT;   313 AA.
AC   P45381;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Aspartoacylase (EC 3.5.1.15) (Aminoacylase-2) (ACY-2).
GN   ASPA OR ASP OR ACY2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND VARIANT CD ALA-285.
RC   TISSUE=Kidney;
RX   MEDLINE=94073185; PubMed=8252036;
RA   Kaul R., Gao G.P., Balamurugan K., Matalon R.;
RT   "Cloning of the human aspartoacylase cDNA and a common missense
RT   mutation in Canavan disease.";
RL   Nat. Genet. 5:118-123(1993).
RN   [2]
RP   VARIANTS CD ALA-285 AND GLU-305.
RX   MEDLINE=94295593; PubMed=8023850;
RA   Kaul R., Gao G.P., Aloya M., Balamurugan K., Petrosky A., Michals K.,
RA   Matalon R.;
RT   "Canavan disease: mutations among Jewish and non-jewish patients.";
RL   Am. J. Hum. Genet. 55:34-41(1994).
RN   [3]
RP   VARIANTS CD 176-GLY-ILE-177 DEL; ARG-274; SER-295 AND GLU-305.
RX   MEDLINE=95397842; PubMed=7668285;
RA   Shaag A., Anikster Y., Christensen E., Glustein J.Z., Fois A.,
RA   Michelakakis H., Nigro F., Pronicka E., Ribes A., Zabot M.-T.,
RA   Elpeleg O.N.;
RT   "The molecular basis of canavan (aspartoacylase deficiency) disease
RT   in European non-Jewish patients.";
RL   Am. J. Hum. Genet. 57:572-580(1995).
RN   [4]
```

```
RP   VARIANT CD ARG-152.
RX   MEDLINE=95322986; PubMed=7599639;
RA   Kaul R., Gao G.P., Michals K., Whelan D.T., Levin S., Matalon R.;
RT   "Novel (Cys152 > Arg) missense mutation in an Arab patient with
RT   Canavan disease.";
RL   Hum. Mutat. 5:269-271(1995).
RN   [5]
RP   VARIANT CD THR-143.
RX   MEDLINE=98112496; PubMed=9452117;
RA   Kobayashi K., Tsujino S., Ezoe T., Hamaguchi H., Nihei K.,
RA   Sakuragawa N.;
RT   "Missense mutation (I143T) in a Japanese patient with Canavan
RT   disease.";
RL   Hum. Mutat. Suppl. 1:S308-S309(1998).
CC   -!- FUNCTION: RESPONSIBLE FOR THE CLEAVAGE OF N-ACETYLASPARTIC ACID
CC       (NAA) WHICH OCCURS IN HIGH CONCENTRATION IN BRAIN. HYDROLYSIS OF
CC       NAA PLAYS A SIGNIFICANT PART IN THE MAINTENANCE OF INTACT WHITE
CC       MATTER. IN OTHER TISSUES ACTS AS A SCAVENGER OF NAA FROM BODY
CC       FLUIDS.
CC   -!- CATALYTIC ACTIVITY: N-acyl-L-aspartate + H(2)O = a fatty acid
CC       anion + L-aspartate.
CC   -!- TISSUE SPECIFICITY: BRAIN WHITE MATTER, SKELETAL MUSCLE, KIDNEY,
CC       ADRENAL GLANDS, LUNG AND LIVER.
CC   -!- DISEASE: Defects in ASPA are the cause of Canavan disease (CD),
CC       also known as spongy degeneration of the brain. CD is an autosomal
CC       recessive leukodystrophy clinically characterized by atonia of
CC       neck muscles, hyperextension of legs and flexion of arms,
CC       blindness, severe mental defect, megalocephaly, and death by 18
CC       months on the average.
CC   -!- SIMILARITY: TO BOVINE ASPARTOACYLASE.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; S67156; AAB29190.1; -.
DR   PIR; S38538; S38538.
DR   Genew; HGNC:756; ASPA.
DR   MIM; 271900; -.
DR   GO; GO:0004046; F:aminoacylase activity; TAS.
DR   GO; GO:0006533; P:aspartate catabolism; TAS.
DR   InterPro; IPR007036; Aste_AspA.
DR   Pfam; PF04952; Aste_AspA; 1.
KW   Hydrolase; Disease mutation.
FT   ACT_SITE     21     21       POTENTIAL.
FT   ACT_SITE     285    285      POTENTIAL.
FT   VARIANT      143    143      I -> T (IN CD; IN A JAPANESE PATIENT).
FT                                /FTId=VAR_004995.
FT   VARIANT      152    152      C -> R (IN CD; LOSS OF ACTIVITY).
FT                                /FTId=VAR_004996.
FT   VARIANT      176    177      MISSING (IN CD).
FT                                /FTId=VAR_004997.
FT   VARIANT      274    274      G -> R (IN CD).
FT                                /FTId=VAR_004998.
FT   VARIANT      285    285      E -> A (IN CD; 2.5% OF ACTIVITY;
FT                                PREDOMINANT MUTATION IN ASHKENAZI JEWISH
FT                                POPULATION).
FT                                /FTId=VAR_004999.
FT   VARIANT      295    295      F -> S (IN CD).
FT                                /FTId=VAR_005000.
FT   VARIANT      305    305      A -> E (IN CD; LOSS OF ACTIVITY; PAN-
FT                                EUROPEAN ORIGIN; MOST PREVALENT AMONG
FT                                NON-JEWISH CD PATIENTS; PROBABLY THE MOST
FT                                ANCIENT MUTATION).
FT                                /FTId=VAR_005001.
SQ   SEQUENCE   313 AA;  35735 MW;  33C0B9B07839E7F5 CRC64;

  Query Match             93.5%;  Score 29;  DB 1;  Length 313;
  Best Local Similarity   83.3%;  Pred. No. 15;

  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              |||:||
Db        136 LIQMFH 141


RESULT 5
YXER_BACSU
ID   YXER_BACSU     STANDARD;      PRT;   370 AA.
AC   P54957;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hypothetical protein yxeR.
GN   YXER OR LP6A.
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168 / BGSC1A1;
RX   MEDLINE=97021444; PubMed=8867804;
RA   Yoshida K.-I., Fujimyra M., Yanai N., Fujita Y.;
RT   "Cloning and sequencing of a 23-kb region of the Bacillus subtilis
RT   genome between the iol and hut operons.";
RL   DNA Res. 2:295-301(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RX   MEDLINE=98044033; PubMed=9384377;
RA   Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA   Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA   Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA   Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA   Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA   Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA   Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA   Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA   Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA   Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA   Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA   Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA   Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
RA   Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA   Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA   Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA   Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA   Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA   Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA   Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA   Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA   Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA   Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA   Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA   Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA   Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,
RA   Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA   Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT   "The complete genome sequence of the Gram-positive bacterium Bacillus
RT   subtilis.";
RL   Nature 390:249-256(1997).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC   -!- SIMILARITY: TO S.TYPHIMURIUM ETHANOLAMINE UTILIZATION PROTEIN
CC       EUTH.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
```

```
CC   ----------------------------------------------------------------------
DR   EMBL; D45912; BAA08334.1; -.
DR   EMBL; Z99124; CAB15981.1; -.
DR   PIR; D70076; D70076.
DR   SubtiList; BG11894; yxeR.
DR   Pfam; PF04346; EutH; 1.
KW   Hypothetical protein; Transmembrane; Transport; Complete proteome.
FT   TRANSMEM      6     26       POTENTIAL.
FT   TRANSMEM     49     69       POTENTIAL.
FT   TRANSMEM     79     99       POTENTIAL.
FT   TRANSMEM    111    131       POTENTIAL.
FT   TRANSMEM    143    163       POTENTIAL.
FT   TRANSMEM    167    187       POTENTIAL.
FT   TRANSMEM    206    226       POTENTIAL.
FT   TRANSMEM    236    256       POTENTIAL.
FT   TRANSMEM    307    327       POTENTIAL.
FT   TRANSMEM    333    353       POTENTIAL.
SQ   SEQUENCE   370 AA;  38414 MW;  E7BF1585EB600DF8 CRC64;

  Query Match             93.5%;  Score 29;  DB 1;  Length 370;
  Best Local Similarity   83.3%;  Pred. No. 18;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              :|||||
Db        193 MIQLFH 198


RESULT 6
FPPS_SCHPO
ID   FPPS_SCHPO     STANDARD;      PRT;   347 AA.
AC   O14230; P78756;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Probable farnesyl pyrophosphate synthetase (FPP synthetase) (FPS)
DE   (Farnesyl diphosphate synthetase) [Includes: Dimethylallyltransferase
DE   (EC 2.5.1.1); Geranyltranstransferase (EC 2.5.1.10)].
GN   SPAC6F12.13C.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RX   MEDLINE=21848401; PubMed=11859360;
RA   Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA   Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA   Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA   Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA   Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA   Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA   James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA   Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA   Oliver K., O'Neil S., Pearson D., Quail M.A., Rabbinowitsch E.,
RA   Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA   Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA   Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA   Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA   Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA   Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
RA   Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA   Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA   Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA   Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA   Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA   Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA   Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA   Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA   Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT   "The genome sequence of Schizosaccharomyces pombe.";
RL   Nature 415:871-880(2002).
RN   [2]
RP   SEQUENCE OF 16-347 FROM N.A.
RC   STRAIN=PR745;
RX   MEDLINE=98162722; PubMed=9501991;
RA   Yoshioka S., Kato K., Nakai K., Okayama H., Nojima H.;
RT   "Identification of open reading frames in Schizosaccharomyces pombe
RT   cDNAs.";
RL   DNA Res. 4:363-369(1997).
CC   -!- FUNCTION: CATALYZES THE SEQUENTIAL CONDENSATION OF ISOPENTENYL
CC       PYROPHOSPHATE WITH THE ALLYLIC PYROPHOSPHATES, DIMETHYLALLYL
CC       PYROPHOSPHATE, AND THEN WITH THE RESULTANT GERANYLPYROPHOSPHATE
CC       TO THE ULTIMATE PRODUCT FARNESYL PYROPHOSPHATE.
CC   -!- CATALYTIC ACTIVITY: Dimethylallyl diphosphate + isopentenyl
CC       diphosphate = diphosphate + geranyl diphosphate.
CC   -!- CATALYTIC ACTIVITY: Geranyl diphosphate + isopentenyl diphosphate
CC       = diphosphate + trans,trans-farnesyl diphosphate.
CC   -!- PATHWAY: ISOPRENE BIOSYNTHESIS, CHOLESTEROL BIOSYNTHESIS.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE FPP/GGPP SYNTHETASES FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; Z98533; CAB11097.1; -.
DR   EMBL; D89104; BAA13767.1; -.
DR   PIR; T11664; T11664.
DR   HSSP; P08836; 1FPS.
DR   GeneDB_SPombe; SPAC6F12.13c; -.
DR   InterPro; IPR000092; Polyprenyl_synt.
DR   Pfam; PF00348; polyprenyl_synt; 1.
DR   PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
DR   PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
KW   Transferase; Isoprene biosynthesis; Cholesterol biosynthesis.
FT   ACT_SITE    184    184       BY SIMILARITY.
FT   CONFLICT    203    203       A -> E (IN REF. 2).
FT   CONFLICT    346    347       NK -> ISKFLVLCF (IN REF. 2).
SQ   SEQUENCE   347 AA;  39516 MW;  483029A5FC15B9F5 CRC64;

  Query Match             90.3%;  Score 28;  DB 1;  Length 347;
  Best Local Similarity   83.3%;  Pred. No. 29;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ||:|||
Db        148 LIELFH 153


RESULT 7
PDS5_YEAST
ID   PDS5_YEAST     STANDARD;      PRT;  1277 AA.
AC   Q04264; Q04780;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Sister chromatid cohesion protein PDS5 (Precocious dissociation of
DE   sisters protein 5).
GN   PDS5 OR YMR076C OR YM9582.01C OR YM9916.15C.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c / AB972;
RX   PubMed=9169872;
RA   Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
RA   Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
```

```
RA   Jagels K., Lye G., Moule S., Odell C., Pearson D., Rajandream M.A.,
RA   Rice P., Skelton J., Walsh S., Whitehead S., Barrell B.G.;
RT   "The nucleotide sequence of Saccharomyces cerevisiae chromosome
RT   XIII.";
RL   Nature 387:90-93(1997).
RN   [2]
RP   FUNCTION, AND SUBCELLULAR LOCATION.
RX   MEDLINE-20517447; PubMed-11062262;
RA   Hartman T., Stead K., Koshland D., Guacci V.;
RT   "Pds5p is an essential chromosomal protein required for both sister
RT   chromatid cohesion and condensation in Saccharomyces cerevisiae.";
RL   J. Cell Biol. 151:613-626(2000).
RN   [3]
RP   FUNCTION.
RX   MEDLINE=21065936; PubMed-11137006;
RA   Panizza S., Tanaka T., Hochwagen A., Eisenhaber F., Nasmyth K.;
RT   "Pds5 cooperates with cohesin in maintaining sister chromatid
RT   cohesion.";
RL   Curr. Biol. 10:1557-1564(2000).
CC   -!- FUNCTION: Essential for the establishment and maintenance of
CC       sister chromatid cohesion at centromere proximal and distal
CC       regions during S phase. Also required for chromosomal
CC       condensation.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z49259; CAA89222.1; -.
DR   EMBL; Z48952; CAA88801.1; -.
DR   PIR; S54451; S54451.
DR   SGD; S0004681; PDS5.
DR   GO; GO:0005708; C:mitotic chromosome; IDA.
DR   GO; GO:0007076; P:mitotic chromosome condensation; IMP.
DR   GO; GO:0007064; P:mitotic sister chromatid cohesion; IMP.
KW   Mitosis; Nuclear protein.
SQ   SEQUENCE   1277 AA;  147040 MW;  9DF40A5274FD9623 CRC64;

  Query Match             90.3%;  Score 28;  DB 1;  Length 1277;
  Best Local Similarity   83.3%;  Pred. No. 1.2e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ||:|||
Db        149 LIELFH 154


RESULT 8
MYG_RABIT
ID   MYG_RABIT      STANDARD;      PRT;   153 AA.
AC   P02170;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Myoglobin.
GN   MB.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=76253773; PubMed=952958;
RA   Romero-Herrera A.E., Lehmann H., Castillo O.;
RT   "The primary structure of the myoglobin of rabbit (Oryctolagus
RT   cuniculus).";
RL   Biochim. Biophys. Acta 439:51-54(1976).
CC   -!- FUNCTION: SERVES AS A RESERVE SUPPLY OF OXYGEN AND FACILITATES
CC       THE MOVEMENT OF OXYGEN WITHIN MUSCLES.
CC   -!- SIMILARITY: Belongs to the globin family.
DR   PIR; A02492; MYRB.
DR   HSSP; P02189; 1MWC.
DR   InterPro; IPR000971; Globin.
DR   InterPro; IPR002335; Myoglobin.
DR   Pfam; PF00042; globin; 1.
DR   PRINTS; PR00613; MYOGLOBIN.
DR   PROSITE; PS01033; GLOBIN; 1.
KW   Heme; Oxygen transport; Transport; Muscle.
FT   METAL        64     64       IRON (HEME DISTAL LIGAND).
FT   METAL        93     93       IRON (HEME PROXIMAL LIGAND).
SQ   SEQUENCE   153 AA;  17089 MW;  D77CE1301269EB1F CRC64;

  Query Match             87.1%;  Score 27;  DB 1;  Length 153;
  Best Local Similarity   83.3%;  Pred. No. 20;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ||:|||
Db         29 LIRLFH 34


RESULT 9
RT22_MOUSE
ID   RT22_MOUSE     STANDARD;      PRT;   359 AA.
AC   Q9CXW2; Q8VDG9; Q9D185;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Mitochondrial 28S ribosomal protein S22 (S22mt) (MRP-S22).
GN   MRPS22 OR RPMS22.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Embryo, and Embryonic head;
RX   PubMed=12466851;
RA   Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA   Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA   Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA   Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA   Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA   Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA   Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA   Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA   Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA   Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA   Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA   Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA   Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA   Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA   Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA   Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA   Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA   Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA   Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA   Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA   Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA   Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA   Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA   Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA   Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA   Birney E., Hayashizaki Y.;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
RN   [2]
RP   SEQUENCE OF 254-359 FROM N.A.
RC   STRAIN=Czech II; TISSUE=Mammary gland;
RX   PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
```

```
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- SUBUNIT: Component of the mitochondrial ribosome small subunit
CC       (28S) which comprises a 12S rRNA and about 30 distinct proteins.
CC   -!- SUBCELLULAR LOCATION: Mitochondrial (By similarity).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AK003836; BAB23027.2; -.
DR   EMBL; AK013925; BAB29061.1; -.
DR   EMBL; BC021882; AAH21882.1; ALT_INIT.
DR   MGD; MGI:1928137; Mrps22.
KW   Ribosomal protein; Mitochondrion.
SQ   SEQUENCE   359 AA;  41192 MW;  C258AD152738F1C7 CRC64;

  Query Match             87.1%;  Score 27;  DB 1;  Length 359;
  Best Local Similarity   66.7%;  Pred. No. 52;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              |:|||:|
Db        299 LVQLYH 304


RESULT 10
RT22_HUMAN
ID   RT22_HUMAN     STANDARD;      PRT;   360 AA.
AC   P82650;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Mitochondrial 28S ribosomal protein S22 (S22mt) (MRP-S22) (GK002).
GN   MRPS22 OR RPMS22 OR C3ORF5.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver cancer;
RA   Li Y., Wu T., Xu S., Ren S., Chen Z., Han Z.;
RT   "A novel gene expressed in human liver cancer tissue.";
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21096975; PubMed=11175783;
RA   Crisponi L., Deiana M., Loi A., Chiappe F., Uda M., Amati P.,
RA   Bisceglia L., Zelante L., Nagaraja R., Porcu S., Ristaldi M.S.,
RA   Marzella R., Rocchi M., Nicolino M., Lienhardt-Roussie A., Nivelon A.,
RA   Verloes A., Schlessinger D., Gasparini P., Bonneau D., Cao A.,
RA   Pilia G.;
RT   "The putative forkhead transcription factor FOXL2 is mutated in
RT   blepharophimosis/ptosis/epicanthus inversus syndrome.";
RL   Nat. Genet. 27:159-166(2001).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Muscle;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [4]
RP   IDENTIFICATION.
RX   MEDLINE=20490686; PubMed=10938081;
RA   Koc E.C., Burkhart W., Blackburn K., Moseley A., Koc H.,
RA   Spremulli L.L.;
RT   "A proteomics approach to the identification of mammalian
RT   mitochondrial small subunit ribosomal proteins.";
RL   J. Biol. Chem. 275:32585-32591(2000).
CC   -!- SUBUNIT: Component of the mitochondrial ribosome small subunit
CC       (28S) which comprises a 12S rRNA and about 30 distinct proteins.
CC   -!- SUBCELLULAR LOCATION: Mitochondrial.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF226045; AAF86945.1; -.
DR   EMBL; AF321613; AAK01406.1; -.
DR   EMBL; BC009296; AAH09296.1; -.
DR   Genew; HGNC:14508; MRPS22.
DR   MIM; 605810; -.
DR   GO; GO:0005761; C:mitochondrial ribosome; NAS.
KW   Ribosomal protein; Mitochondrion.
SQ   SEQUENCE   360 AA;  41280 MW;  2DB50257C222D706 CRC64;

  Query Match             87.1%;  Score 27;  DB 1;  Length 360;
  Best Local Similarity   66.7%;  Pred. No. 52;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              |:|||:|
Db        300 LVQLYH 305


RESULT 11
PROA_LEPIN
ID   PROA_LEPIN     STANDARD;      PRT;   416 AA.
AC   P94872;
DT   30-MAY-2000 (Rel. 39, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
```

```
DE   Gamma-glutamyl phosphate reductase (GPR) (EC 1.2.1.41) (Glutamate-5-
DE   semialdehyde dehydrogenase) (Glutamyl-gamma-semialdehyde
DE   dehydrogenase) (GSA dehydrogenase).
GN   PROA OR LA0854.
OS   Leptospira interrogans.
OC   Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX   NCBI_TaxID=173;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Serogroup Icterohaemorrhagiae;
RA   Stamm L.V., Barnes N.Y.;
RL   Submitted (OCT-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RX   MEDLINE=22598143; PubMed=12712204;
RA   Ren S.-X., Fu G., Jiang X.-G., Zeng R., Miao Y.-G., Xu H.,
RA   Zhang Y.-X., Xiong H., Lu G., Lu L.-F., Jiang H.-Q., Jia J., Tu Y.-F.,
RA   Jiang J.-X., Gu W.-Y., Zhang Y.-Q., Cai Z., Sheng H.-H., Yin H.-F.,
RA   Zhang Y., Zhu G.-F., Wan M., Huang H.-L., Qian Z., Wang S.-Y., Ma W.,
RA   Yao Z.-J., Shen Y., Qiang B.-Q., Xia Q.-C., Guo X.-K., Danchin A.,
RA   Saint Girons I., Somerville R.L., Wen Y.-M., Shi M.-H., Chen Z.,
RA   Xu J.-G., Zhao G.-P.;
RT   "Unique physiological and pathogenic features of Leptospira
RT   interrogans revealed by whole-genome sequencing.";
RL   Nature 422:888-893(2003).
CC   -!- FUNCTION: CATALYZES THE NADPH DEPENDENT REDUCTION OF L-GAMMA-
CC       GLUTAMYL 5-PHOSPHATE INTO L-GLUTAMATE 5-SEMIALDEHYDE AND
CC       PHOSPHATE. THE PRODUCT SPONTANEOUSLY UNDERGOES CYCLIZATION TO FORM
CC       1-PYRROLINE-5-CARBOXYLATE.
CC   -!- CATALYTIC ACTIVITY: L-glutamate 5-semialdehyde + phosphate +
CC       NADP(+) = L-gamma-glutamyl 5-phosphate + NADPH.
CC   -!- PATHWAY: Proline biosynthesis; second step.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE GAMMA-GLUTAMYL PHOSPHATE REDUCTASE
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U73651; AAB39854.1; -.
DR   EMBL; AE011270; AAN48053.1; -.
DR   HAMAP; MF_00412; -; 1.
DR   InterPro; IPR002086; Aldehyde_dehydr.
DR   InterPro; IPR000965; Glut_PP_reduct.
DR   Pfam; PF00171; aldedh; 1.
DR   TIGRFAMs; TIGR00407; proA; 1.
DR   PROSITE; PS01223; PROA; FALSE_NEG.
KW   Oxidoreductase; Proline biosynthesis; NADP; Complete proteome.
FT   CONFLICT     32     32       V -> I (IN REF. 1).
FT   CONFLICT    299    299       N -> K (IN REF. 1).
FT   CONFLICT    305    305       E -> Q (IN REF. 1).
SQ   SEQUENCE   416 AA;  45768 MW;  A949288589EB0B9A CRC64;

  Query Match             87.1%;  Score 27;  DB 1;  Length 416;
  Best Local Similarity   83.3%;  Pred. No. 61;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ||:|||
Db        160 LIKLFH 165


RESULT 12
T23E_BACTF
ID   T23E_BACTF     STANDARD;      PRT;   478 AA.
AC   Q02403;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Transposase for insertion sequence element IS231E.
OS   Bacillus thuringiensis (subsp. finitimus).
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=29337;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Serotype 2;
RX   MEDLINE=92374847; PubMed=1324390;
RA   Rezsoehazy R., Hallet B., Delcour J.;
RT   "IS231D, E and F, three new insertion sequences in Bacillus
RT   thuringiensis: extension of the IS231 family.";
RL   Mol. Microbiol. 6:1959-1967(1992).
CC   -!- FUNCTION: INVOLVED IN THE TRANSPOSITION OF THE INSERTION
CC       SEQUENCE.
CC   -!- SIMILARITY: BELONGS TO THE TRANSPOSASE FAMILY 11.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X63384; CAA44990.1; -.
DR   PIR; S25821; S25821.
DR   InterPro; IPR002559; Transposase_11.
DR   Pfam; PF01609; Transposase_11; 1.
KW   Transposable element; Transposition; DNA-binding; DNA recombination.
SQ   SEQUENCE   478 AA;  56087 MW;  A0AAA4F46E61C223 CRC64;

  Query Match             87.1%;  Score 27;  DB 1;  Length 478;
  Best Local Similarity   83.3%;  Pred. No. 71;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ||:|||
Db        436 LIRLFH 441


RESULT 13
HEP2_HUMAN
ID   HEP2_HUMAN     STANDARD;      PRT;   499 AA.
AC   P05546;
DT   01-NOV-1988 (Rel. 09, Created)
DT   01-NOV-1991 (Rel. 20, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Heparin cofactor II precursor (HC-II) (Protease inhibitor leuserpin 2)
DE   (HLS2).
GN   SERPIND1 OR HCF2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91120782; PubMed=1671335;
RA   Herzog R., Lutz S., Blin N., Marasa J.C., Blinder M.A.,
RA   Tollefsen D.M.;
RT   "Complete nucleotide sequence of the gene for human heparin cofactor
RT   II and mapping to chromosomal band 22q11.";
RL   Biochemistry 30:1350-1357(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88163663; PubMed=2894851;
RA   Blinder M.A., Marasa J.C., Reynolds C.H., Deaven L.L., Tollefsen D.M.;
RT   "Heparin cofactor II: cDNA sequence, chromosome localization,
RT   restriction fragment length polymorphism, and expression in
RT   Escherichia coli.";
RL   Biochemistry 27:752-759(1988).
RN   [3]
```

```
RP   SEQUENCE OF 19-499 FROM N.A.
RX   MEDLINE=86120356; PubMed=3003690;
RA   Ragg H.;
RT   "A new member of the plasma protease inhibitor gene family.";
RL   Nucleic Acids Res. 14:1073-1088(1986).
RN   [4]
RP   SEQUENCE OF 333-499 FROM N.A.
RX   MEDLINE=86242236; PubMed=3755044;
RA   Inhorn R.C., Tollefsen D.M.;
RT   "Isolation and characterization of a partial cDNA clone for heparin
RT   cofactor II1.";
RL   Biochem. Biophys. Res. Commun. 137:431-436(1986).
RN   [5]
RP   SEQUENCE OF 20-52 AND 464-499.
RX   MEDLINE=86077723; PubMed=3907702;
RA   Griffith M.J., Noyes C.M., Tyndall J.A., Church F.C.;
RT   "Structural evidence for leucine at the reactive site of heparin
RT   cofactor II.";
RL   Biochemistry 24:6777-6782(1985).
RN   [6]
RP   SEQUENCE OF 1-119 FROM N.A.
RX   MEDLINE=88298901; PubMed=2841345;
RA   Ragg H., Preibisch G.;
RT   "Structure and expression of the gene coding for the human serpin
RT   hLS2.";
RL   J. Biol. Chem. 263:12129-12134(1988).
RN   [7]
RP   SEQUENCE OF 58-85.
RX   MEDLINE=91093260; PubMed=1985958;
RA   Church F.C., Pratt C.W., Hoffman M.;
RT   "Leukocyte chemoattractant peptides from the serpin heparin cofactor
RT   II.";
RL   J. Biol. Chem. 266:704-709(1991).
RN   [8]
RP   FUNCTION OF N-TERMINAL ACIDIC DOMAIN.
RX   MEDLINE=92041850; PubMed=1939083;
RA   van Deerlin V.M.D., Tollefsen D.M.;
RT   "The N-terminal acidic domain of heparin cofactor II mediates the
RT   inhibition of alpha-thrombin in the presence of glycosaminoglycans.";
RL   J. Biol. Chem. 266:20223-20231(1991).
RN   [9]
RP   MUTAGENESIS OF ARG-122 AND LYS-204.
RX   MEDLINE=90094412; PubMed=2104620;
RA   Blinder M.A., Tollefsen D.M.;
RT   "Site-directed mutagenesis of arginine 103 and lysine 185 in the
RT   proposed glycosaminoglycan-binding site of heparin cofactor II.";
RL   J. Biol. Chem. 265:286-291(1990).
RN   [10]
RP   VARIANT OSLO HIS-208.
RX   MEDLINE=89174798; PubMed=2647747;
RA   Blinder M.A., Andersson T.R., Abildgaard U., Tollefsen D.M.;
RT   "Heparin cofactor IIOslo. Mutation of Arg-189 to His decreases the
RT   affinity for dermatan sulfate.";
RL   J. Biol. Chem. 264:5128-5133(1989).
RN   [11]
RP   VARIANTS THR-7; HIS-208 AND MET-442.
RX   MEDLINE=99318093; PubMed=10391209;
RA   Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
RA   Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
RA   Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,
RA   Lander E.S.;
RT   "Characterization of single-nucleotide polymorphisms in coding regions
RT   of human genes.";
RL   Nat. Genet. 22:231-238(1999).
RN   [12]
RP   ERRATUM.
RA   Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
RA   Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
RA   Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,
RA   Lander E.S.;
RL   Nat. Genet. 23:373-373(1999).
CC   -!- FUNCTION: THROMBIN INHIBITOR ACTIVATED BY THE GLYCOSAMINOGLYCANS,
CC       HEPARIN OR DERMATAN SULFATE. IN THE PRESENCE OF THE LATTER,
CC       HC-II BECOMES THE PREDOMINANT THROMBIN INHIBITOR IN PLACE OF
CC       ANTITHROMBIN III (AT). ALSO INHIBITS CHYMOTRYPSIN, BUT IN A
CC       GLYCOSAMINOGLYCAN-INDEPENDENT MANNER.
CC   -!- FUNCTION: PEPTIDES AT THE N-TERMINAL OF HC-II HAVE CHEMOTACTIC
CC       ACTIVITY FOR BOTH MONOCYTES AND NEUTROPHILS.
CC   -!- TISSUE SPECIFICITY: Expressed predominantly in liver.
CC   -!- DOMAIN: THE N-TERMINAL ACIDIC REPEAT REGION MEDIATES, IN PART, THE
CC       GLYCOSAMINOGLYCAN-ACCELERATED THROMBIN INHIBITION.
CC   -!- DISEASE: DEFECTS IN SERPIND1 ARE ASSOCIATED WITH THROMBOSIS
CC       (THROMBOPHILIA).
CC   -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; M58600; AAA52641.1; -.
DR   EMBL; X03498; CAA27218.1; ALT_INIT.
DR   EMBL; M12849; AAA52642.1; -.
DR   EMBL; M33660; AAA36185.1; -.
DR   PIR; A37924; A37924.
DR   PDB; 1JMJ; 30-AUG-02.
DR   PDB; 1JMO; 30-AUG-02.
DR   Genew; HGNC:4838; SERPIND1.
DR   MIM; 142360; -.
DR   InterPro; IPR000295; Leuserpin2.
DR   InterPro; IPR000215; Serpin.
DR   Pfam; PF00079; serpin; 1.
DR   PRINTS; PR00780; LEUSERPINII.
DR   SMART; SM00093; SERPIN; 1.
DR   PROSITE; PS00284; SERPIN; 1.
KW   Serpin; Serine protease inhibitor; Glycoprotein; Heparin-binding;
KW   Plasma; Blood coagulation; Chemotaxis; Signal; Thrombophilia;
KW   Disease mutation; Sulfation; Repeat; Polymorphism; 3D-structure.
FT   SIGNAL       1    19
FT   CHAIN       20   499       HEPARIN COFACTOR II.
FT   DOMAIN      68    79       CHEMOTACTIC ACTIVITY.
FT   DOMAIN      73    97       2 X 11 AA APPROXIMATE REPEATS, ASP/GLU-
FT                              RICH (ACIDIC) (HIRUDIN-LIKE).
FT   REPEAT      73    83       1.
FT   REPEAT      87    97       2.
FT   DOMAIN     192   212       GLYCOSAMINOGLYCAN-BINDING SITE.
FT   MOD_RES     79    79       SULFATION.
FT   MOD_RES     92    92       SULFATION.
FT   CARBOHYD    49    49       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   188   188       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   387   387       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   ACT_SITE   463   464       REACTIVE BOND (BY SIMILARITY).
FT   VARIANT      7     7       A -> T (IN dbSNP:5905).
FT                              /FTId=VAR_011746.
FT   VARIANT     60    60       H -> P (IN dbSNP:165867).
FT                              /FTId=VAR_011747.
FT   VARIANT    208   208       R -> H (IN OSLO; HCF2 DEFICIENCY;
FT                              DECREASED AFFINITY FOR DERMATAN SULFATE;
FT                              dbSNP:6065).
FT                              /FTId=VAR_007112.
FT   VARIANT    237   237       K -> R (IN dbSNP:1042435).
FT                              /FTId=VAR_011748.
FT   VARIANT    442   442       T -> M (IN dbSNP:5904).
FT                              /FTId=VAR_011749.
FT   MUTAGEN    122   122       R->L: NORMAL THROMBIN INHIBITION AND
FT                              GLYCOSAMINOGLYCAN AFFINITY.
FT   MUTAGEN    122   122       R->Q: GREATLY REDUCED THROMBIN
FT                              INHIBITION. NORMAL GLYCOSAMINOGLYCAN
FT                              AFFINITY.
FT   MUTAGEN    122   122       R->W: GREATLY REDUCED THROMBIN
FT                              INHIBITION. NORMAL GLYCOSAMINOGLYCAN
FT                              AFFINITY.
FT   MUTAGEN    204   204       K->M: REDUCED HEPARIN- AND NO DERMATAN
```

```
FT                                   SULFATE-ACTIVATED INHIBITION.
FT   MUTAGEN      204    204         K->N: REDUCED HEPARIN- AND NO DERMATAN
FT                                   SULFATE-ACTIVATED INHIBITION.
FT   MUTAGEN      204    204         K->T: REDUCED HEPARIN- AND NO DERMATAN
FT                                   SULFATE-ACTIVATED INHIBITION.
FT   CONFLICT      49     49         MISSING (IN REF. 5).
FT   CONFLICT     483    483         R -> P (IN REF. 5).
FT   CONFLICT     486    486         C -> T (IN REF. 5).
FT   CONFLICT     499    499         S -> Q (IN REF. 5).
SQ   SEQUENCE   499 AA;  57070 MW;  3BCE353FE1F6DF05 CRC64;

  Query Match             87.1%;  Score 27;  DB 1;  Length 499;
  Best Local Similarity   66.7%;  Pred. No. 74;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ::||||
Db        113 ILQLFH 118


RESULT 14
YNA8_YEAST
ID   YNA8_YEAST     STANDARD;      PRT:   669 AA.
AC   P53983;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   Hypothetical 76.7 kDa protein in SPO1-SIS1 intergenic region.
GN   YNL008C OR N2874.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE OF 1-102 FROM N.A.
RA   Andre B., Iraqui Houssaini I., Urrestarazu L.A., Vissers S.;
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 371-669 FROM N.A.
RA   Doignon F., Crouzet M.;
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z71284; CAA95868.1; -.
DR   EMBL; Z71283; CAA95867.1; -.
DR   PIR; S62919; S62919.
DR   SGD; S0004953; ASI3.
DR   GO; GO:0005789; C:endoplasmic reticulum membrane; IGI.
DR   GO; GO:0004842; F:ubiquitin-protein ligase activity; IGI.
DR   GO; GO:0006511; P:ubiquitin-dependent protein catabolism; IGI.
KW   Hypothetical protein; Transmembrane.
FT   TRANSMEM      72     92        POTENTIAL.
FT   TRANSMEM     120    140        POTENTIAL.
FT   TRANSMEM     150    170        POTENTIAL.
FT   TRANSMEM     175    195        POTENTIAL.
FT   TRANSMEM     209    229        POTENTIAL.
FT   TRANSMEM     234    254        POTENTIAL.
FT   TRANSMEM     271    291        POTENTIAL.
SQ   SEQUENCE   669 AA;  76741 MW;  1BB4DCAC74E8FF63 CRC64;

  Query Match             87.1%;  Score 27;  DB 1;  Length 669;
  Best Local Similarity   66.7%;  Pred. No. 1e+02;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ::||||
Db        223 LVELFH 228


RESULT 15
CLC6_HUMAN
ID   CLC6_HUMAN     STANDARD;      PRT:   869 AA.
AC   P51797; O60818; O60819; O60820; O60821; P78520; P78521; Q99427;
AC   Q99428; Q99429;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Chloride channel protein 6 (ClC-6).
GN   CLCN6 OR KIAA0046.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM A).
RC   TISSUE=Brain;
RX   MEDLINE=96130311; PubMed=8543009;
RA   Brandt S., Jentsch T.J.;
RT   "ClC-6 and ClC-7 are two novel broadly expressed members of the CLC
RT   chloride channel family.";
RL   FEBS Lett. 377:15-20(1995).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORMS B; C AND D), AND SEQUENCE OF 1-409 FROM
RP   N.A. (ISOFORM A).
RC   TISSUE=Chronic myeloid leukemia cell;
RX   MEDLINE=97344267; PubMed=9224655;
RA   Eggermont J., Buyse G., Voets T., Tytgat J., De Smedt H.,
RA   Droogmans G., Nilius B.;
RT   "Alternative splicing of ClC-6 (a member of the ClC chloride-channel
RT   family) transcripts generates three truncated isoforms one of which,
RT   ClC-6c, is kidney-specific.";
RL   Biochem. J. 325:269-276(1997).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM A).
RX   MEDLINE=99431565; PubMed=10500249;
RA   Kornak U., Boesl M.R., Kubisch C.;
RT   "Complete genomic structure of the CLCN6 and CLCN7 putative chloride
RT   channel genes.";
RL   Biochim. Biophys. Acta 1447:100-106(1999).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORM A).
RC   TISSUE=Brain;
RA   Ohara O., Nagase T., Kikuno R., Nomura N.;
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE OF 50-869 FROM N.A., AND ALTERNATIVE SPLICING.
RA   Errington H.;
RL   Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE OF 237-869 FROM N.A. (ISOFORM A).
RC   TISSUE=Bone marrow;
RX   MEDLINE=96051398; PubMed=7584044;
RA   Nomura N., Nagase T., Miyajima N., Sazuka T., Tanaka A., Sato S.,
RA   Seki N., Kawarabayasi Y., Ishikawa K.-I., Tabata S.;
RT   "Prediction of the coding sequences of unidentified human genes. II.
RT   The coding sequences of 40 new genes (KIAA0041-KIAA0080) deduced by
RT   analysis of cDNA clones from human cell line KG-1.";
RL   DNA Res. 1:223-229(1994).
RN   [7]
RP   TISSUE SPECIFICITY.
RC   TISSUE=Aortic endothelium, and Vascular smooth muscle;
RX   MEDLINE=99222497; PubMed=10198195;
RA   Lamb F.S., Clayton G.H., Liu B.-X., Smith R.L., Barna T.J.,
RA   Schutte B.C.;
RT   "Expression of CLCN voltage-gated chloride channel genes in human
RT   blood vessels.";
RL   J. Mol. Cell. Cardiol. 31:657-666(1999).
CC   -!- FUNCTION: VOLTAGE-GATED CHLORIDE CHANNEL. CHLORIDE CHANNELS HAVE
```

```
CC       SEVERAL FUNCTIONS INCLUDING THE REGULATION OF CELL VOLUME;
CC       MEMBRANE POTENTIAL STABILIZATION, SIGNAL TRANSDUCTION AND
CC       TRANSEPITHELIAL TRANSPORT.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=4;
CC       Name=A; Synonyms=ClC-6a;
CC         IsoId=P51797-1; Sequence=Displayed;
CC       Name=B; Synonyms=ClC-6b, D2-A1;
CC         IsoId=P51797-2; Sequence=VSP_001043, VSP_001044;
CC       Name=C; Synonyms=ClC-6c, D1-A1;
CC         IsoId=P51797-3; Sequence=VSP_001045, VSP_001046;
CC       Name=D; Synonyms=ClC-6d, D1-A2;
CC         IsoId=P51797-4; Sequence=VSP_001047, VSP_001048;
CC   -!- TISSUE SPECIFICITY: Testis, ovary, small intestine, brain and
CC       skeletal muscle. Low level expression in aortic and coronary
CC       vascular smooth muscle cells, and aortic endothelial cells.
CC       Isoform C is only detected in kidney.
CC   -!- SIMILARITY: Belongs to the chloride channel family.
CC   -!- SIMILARITY: Contains 2 CBS domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X83378; CAA58292.1; -.
DR   EMBL; X96391; CAA65255.1; -.
DR   EMBL; X99472; CAA67835.1; -.
DR   EMBL; X99473; CAA67836.1; -.
DR   EMBL; X99474; CAA67837.1; -.
DR   EMBL; X99475; CAA67838.1; -.
DR   EMBL; AF009257; AAB69287.1; -.
DR   EMBL; AF009247; AAB69287.1; JOINED.
DR   EMBL; AF009248; AAB69287.1; JOINED.
DR   EMBL; AF009249; AAB69287.1; JOINED.
DR   EMBL; AF009250; AAB69287.1; JOINED.
DR   EMBL; AF009251; AAB69287.1; JOINED.
DR   EMBL; AF009252; AAB69287.1; JOINED.
DR   EMBL; AF009253; AAB69287.1; JOINED.
DR   EMBL; AF009254; AAB69287.1; JOINED.
DR   EMBL; AF009255; AAB69287.1; JOINED.
DR   EMBL; AF009256; AAB69287.1; JOINED.
DR   EMBL; AL021155; CAA15951.1; -.
DR   EMBL; AL021155; CAA15952.1; -.
DR   EMBL; AL021155; CAA15953.1; -.
DR   EMBL; AL021155; CAA15954.1; -.
DR   EMBL; D28475; BAA05836.3; -.
DR   Genew; HGNC:2024; CLCN6.
DR   MIM; 602726; -.
DR   GO; GO:0005887; C:integral to plasma membrane; TAS.
DR   GO; GO:0005247; F:voltage-gated chloride channel activity; NAS.
DR   GO; GO:0006821; P:chloride transport; NAS.
DR   GO; GO:0006884; P:regulation of cell volume; NAS.
DR   GO; GO:0007165; P:signal transduction; NAS.
DR   GO; GO:0006832; P:small molecule transport; TAS.
DR   InterPro; IPR000644; CBS_domain.
DR   InterPro; IPR001807; Cl-channel_volt.
DR   Pfam; PF00571; CBS; 2.
DR   Pfam; PF00654; voltage_CLC; 1.
DR   PRINTS; PR00762; CLCHANNEL.
KW   Ionic channel; Ion transport; Chloride channel; Chloride;
KW   Voltage-gated channel; Transmembrane; CBS domain; Repeat;
KW   Alternative splicing.
FT   TRANSMEM     82    102       POTENTIAL.
FT   TRANSMEM    130    150       POTENTIAL.
FT   TRANSMEM    184    204       POTENTIAL.
FT   TRANSMEM    245    265       POTENTIAL.
FT   TRANSMEM    334    354       POTENTIAL.
FT   TRANSMEM    375    395       POTENTIAL.
FT   TRANSMEM    462    482       POTENTIAL.
FT   TRANSMEM    511    531       POTENTIAL.
FT   DOMAIN      603    658       CBS 1.
FT   DOMAIN      805    856       CBS 2.
FT   DOMAIN        4     18       CYS-RICH.
FT   VARSPLIC    237    320       DKRDFVSAGAAAGVAAAFGAPIGGTLFSLEEGSSFWNQGLT
FT                                WKVLFCSMSATFTLNFFRSGIQFGSWGSFQLPGLLNFGEFK
FT                                CS -> YGKRQERLCISRSGCWSCCSFRGANRGYLVQSRGG
FT                                FVLLEPRAHVESALLFHVCHLHPQLLPFWDSVWKLGFLPAP
FT                                WIAELWRV (in isoform B).
FT                                /FTId=VSP_001043.
FT   VARSPLIC    321    869       Missing (in isoform B).
FT                                /FTId=VSP_001044.
FT   VARSPLIC    319    353       CSDSDKKCHLWTAMDLGFFVVMGVIGGLLGATFNC -> SL
FT                                REPPCVSGNHRGGVCGLDGVRRMPTDVLFESNR (in
FT                                isoform C).
FT                                /FTId=VSP_001045.
FT   VARSPLIC    354    869       Missing (in isoform C).
FT                                /FTId=VSP_001046.
FT   VARSPLIC    237    308       DKRDFVSAGAAAGVAAAFGAPIGGTLFSLEEGSSFWNQGLT
FT                                WKVLFCSMSATFTLNFFRSGIQFGSWGSFQL -> SGCWSC
FT                                CSFRGANRGYLVQSRGGFVLLEPRAHVESALLFHVCHLHPQ
FT                                LLPFWDSVWKLGFLPAPWIAELWRV (in isoform D).
FT                                /FTId=VSP_001047.
FT   VARSPLIC    309    869       Missing (in isoform D).
FT                                /FTId=VSP_001048.
FT   CONFLICT    198    198       G -> E (IN REF. 5).
SQ   SEQUENCE   869 AA;  97216 MW;  E4273777D9FF4328 CRC64;

  Query Match             87.1%;  Score 27;  DB 1;  Length 869;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+02;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ::||||
Db        450 ILQLFH 455


Search completed: October 14, 2003, 14:34:03
Job time : 6.41667 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 14:33:18 ; Search time 19.875 Seconds
(without alignments)
77.903 Million cell updates/sec

Title: US-09-881-490-150
Perfect score: 31
Sequence: 1 LIQLFH 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	31	100.0	146	16	Q9CHG1	Q9chg1 lactococcus
2	31	100.0	487	4	Q8IW58	Q8iw58 homo sapien
3	31	100.0	898	16	Q8DP20	Q8dp20 streptococc
4	31	100.0	914	16	Q97PQ2	Q97pq2 streptococc
5	30	96.8	1039	2	P72534	P72534 streptococc
6	30	96.8	3152	12	Q8UZ37	Q8uz37 leek yellow
7	29	93.5	136	11	Q8BZC2	Q8bzc2 mus musculu
8	29	93.5	312	11	Q9JIN8	Q9jin8 mus musculu
9	29	93.5	312	11	Q9DCE2	Q9dce2 mus musculu
10	29	93.5	312	11	Q8R3P0	Q8r3p0 mus musculu
11	29	93.5	312	11	Q9D876	Q9d876 mus musculu
12	29	93.5	314	11	Q9R1T5	Q9r1t5 rattus norv
13	29	93.5	379	5	Q9VX89	Q9vx89 drosophila
14	29	93.5	1627	10	Q8LQJ0	Q8lqj0 oryza sativ
15	28	90.3	179	16	Q9KGH7	Q9kgh7 bacillus ha
16	28	90.3	212	3	Q92334	Q92334 sphaceloma
17	28	90.3	261	2	Q9AQB6	Q9aqb6 actinobacil
18	28	90.3	288	5	Q8IAB2	Q8iab2 caenorhabdi
19	28	90.3	351	16	Q8RGG6	Q8rgg6 fusobacteri
20	28	90.3	469	4	Q9NPT5	Q9npt5 homo sapien
21	28	90.3	659	4	O43764	O43764 homo sapien
22	28	90.3	697	4	Q9NXJ9	Q9nxj9 homo sapien
23	28	90.3	703	11	Q91YI9	Q91yi9 mus musculu
24	28	90.3	706	4	Q8NB85	Q8nb85 homo sapien
25	28	90.3	711	5	Q18098	Q18098 caenorhabdi
26	28	90.3	717	4	Q8IUN6	Q8iun6 homo sapien
27	28	90.3	721	4	Q96KL1	Q96kl1 homo sapien
28	28	90.3	721	11	Q9QWV2	Q9qwv2 mus musculu
29	28	90.3	723	4	Q8N1B4	Q8n1b4 homo sapien
30	28	90.3	723	11	O55166	O55166 rattus norv
31	28	90.3	723	11	Q8C754	Q8c754 mus musculu
32	28	90.3	723	11	Q8BQ15	Q8bq15 mus musculu
33	28	90.3	833	11	Q923L0	Q92310 mus musculu
34	28	90.3	1095	4	Q9HCG6	Q9hcg6 homo sapien
35	28	90.3	1904	13	Q8AV28	Q8av28 gallus gall
36	28	90.3	2003	11	Q91Y51	Q91y51 mus musculu
37	28	90.3	2024	4	Q15154	Q15154 homo sapien
38	28	90.3	2025	11	Q9R0L6	Q9r016 mus musculu
39	28	90.3	2031	13	Q9PVV4	Q9pvv4 xenopus lae
40	28	90.3	2075	13	Q90WA4	Q90wa4 fugu rubrip
41	28	90.3	2075	13	Q8JG44	Q8jq44 fugu rubrip
42	28	90.3	2301	10	O80784	O80784 arabidopsis
43	27	87.1	24	5	Q94371	Q94371 caenorhabdi
44	27	87.1	110	11	Q8VDG9	Q8vdg9 mus musculu
45	27	87.1	143	2	Q9Z5V2	Q9z5v2 bacillus th

ALIGNMENTS

RESULT 1
Q9CHG1
ID Q9CHG1 PRELIMINARY; PRT; 146 AA.
AC Q9CHG1;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Transcriptional regulator.
GN RMAG OR LL0770.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IL1403;
RX MEDLINE=21235186; PubMed=11337471;
RA Bolotin A., Wincker P., Mauger S., Jaillon O., Malarme K.,
RA Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis ssp. lactis IL1403.";
RL Genome Res. 11:731-753(2001).
CC -!- SIMILARITY: BELONGS TO THE MARR FAMILY OF TRANSCRIPTIONAL
CC REGULAIORS.
DR EMBL; AE006310; AAK04868.1; -.
DR InterPro; IPR000835; HTH_MarR.
DR Pfam; PF01047; MarR; 1.
DR PRINTS; PR00598; HTHMARR.
DR SMART; SM00347; HTH_MARR; 1.
KW DNA-binding; Transcription regulation; Complete proteome.
SQ SEQUENCE 146 AA; 16808 MW; 30CF04E38EB45953 CRC64;

Query Match 100.0%; Score 31; DB 16; Length 146;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LIQLFH 6
||||||
Db 12 LIQLFH 17

```
RESULT 2
Q8IW58
ID   Q8IW58      PRELIMINARY;      PRT;   487 AA.
AC   Q8IW58;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Similar to bactericidal/permeability-increasing protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Blood;
RA   Strausberg R.;
RL   Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC040955; AAH40955.1; -.
SQ   SEQUENCE   487 AA;  53880 MW;  FE709D9317E5206D CRC64;

  Query Match             100.0%;  Score 31;  DB 4;  Length 487;
  Best Local Similarity   100.0%;  Pred. No. 48;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ||||||
Db        185 LIQLFH 190


RESULT 3
Q8DP20
ID   Q8DP20      PRELIMINARY;      PRT;   898 AA.
AC   Q8DP20;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   P-type ATPase-calcium transporter (EC 3.6.1.36).
GN   PACL OR SPR1410.
OS   Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=171101;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21429245; PubMed=11544234;
RA   Hoskins J., Alborn W.E. Jr., Arnold J., Blaszczak L.C., Burgett S.,
RA   DeHoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,
RA   Gilmour R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E.,
RA   LeBlanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,
RA   McAhren S.M., McHenney M., McLeaster K., Mundy C.W., Nicas T.I.,
RA   Norris F.H., O'Gara M., Peery R.B., Robertson G.T., Rockey P.,
RA   Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,
RA   Zook C.A., Baltz R.H., Jaskunas S.R., Rosteck P.R. Jr., Skatrud P.L.,
RA   Glass J.I.;
RT   "Genome of the bacterium Streptococcus pneumoniae strain R6.";
RL   J. Bacteriol. 183:5709-5717(2001).
DR   EMBL; AE008510; AAL00214.1; -.
KW   Hydrolase; Complete proteome.
SQ   SEQUENCE   898 AA;  97039 MW;  A7AF2CDFBBF9E6D4 CRC64;

  Query Match             100.0%;  Score 31;  DB 16;  Length 898;
  Best Local Similarity   100.0%;  Pred. No. 85;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ||||||
Db        806 LIQLFH 811


RESULT 4
Q97PQ2
ID   Q97PQ2      PRELIMINARY;      PRT;   914 AA.
AC   Q97PQ2;
DT   01-OCT-2001 (TrEMBLrel. 18, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Cation-transporting ATPase, E1-E2 family.
GN   SP1551.
OS   Streptococcus pneumoniae.
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1313;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TIGR4;
RX   MEDLINE=21357209; PubMed=11463916;
RA   Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA   Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,
RA   Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA   Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA   Holtzapple E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,
RA   McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,
RA   Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA   Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT   "Complete genome sequence of a virulent isolate of Streptococcus
RT   pneumoniae.";
RL   Science 293:498-506(2001).
DR   EMBL; AE007449; AAK75638.1; -.
DR   TIGR; SP1551; -.
DR   InterPro; IPR001757; ATPase_E1-E2.
DR   InterPro; IPR006068; Cation_ATPase_C.
DR   InterPro; IPR004014; Cation_ATPase_N.
DR   InterPro; IPR005834; Hydrolase.
DR   InterPro; IPR000695; H_ATPase.
DR   Pfam; PF00689; Cation_ATPase_C; 1.
DR   Pfam; PF00690; Cation_ATPase_N; 1.
DR   Pfam; PF00122; E1-E2_ATPase; 1.
DR   Pfam; PF00702; Hydrolase; 1.
DR   PRINTS; PR00119; CATATPASE.
DR   PRINTS; PR00120; HATPASE.
DR   TIGRFAMs; TIGR01494; ATPase_P-type; 6.
DR   PROSITE; PS00154; ATPASE_E1_E2; 1.
KW   Complete proteome.
SQ   SEQUENCE   914 AA;  99051 MW;  C86323039FE51850 CRC64;

  Query Match             100.0%;  Score 31;  DB 16;  Length 914;
  Best Local Similarity   100.0%;  Pred. No. 86;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ||||||
Db        822 LIQLFH 827


RESULT 5
P72534
ID   P72534      PRELIMINARY;      PRT;  1039 AA.
AC   P72534;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Protein F2.
GN   PRTF2.
OS   Streptococcus pyogenes.
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1314;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=100076;
RX   MEDLINE=97011581; PubMed=8858591;
RA   Jaffe J., Natanson-Yaron S., Caparon M.G., Hanski E.;
RT   "Protein F2, a novel fibronectin-binding protein from Streptococcus
RT   pyogenes, possesses two binding domains.";
RL   Mol. Microbiol. 21:373-384(1996).
```

```
DR   EMBL; U31980; AAC44522.1; -.
DR   InterPro; IPR004237; Fn_bind.
DR   Pfam; PF02986; Fn_bind; 4.
SQ   SEQUENCE   1039 AA;  115037 MW;  DEC806A161A2EC3A CRC64;

  Query Match             96.8%;  Score 30;  DB 2;  Length 1039;
  Best Local Similarity   83.3%;  Pred. No. 1.7e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              |:||||
Db        972 LVQLFH 977


RESULT 6
Q8UZ37
ID   Q8UZ37      PRELIMINARY;      PRT;  3152 AA.
AC   Q8UZ37;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Polyprotein.
OS   Leek yellow stripe potyvirus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC   Potyvirus.
OX   NCBI_TaxID=42004;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Yuhang GYH;
RX   MEDLINE=21887390; PubMed=11890533;
RA   Chen J., Chen J.P., Adams M.J.;
RT   "Characterisation of some carla- and potyviruses from bulb crops in
RT   China.";
RL   Arch. Virol. 147:419-428(2002).
CC   -!- SIMILARITY: BELONGS TO THE POTYVIRUSES POLYPROTEIN FAMILY.
DR   EMBL; AJ307057; CAC85226.1; -.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR001650; Helicase_C.
DR   InterPro; IPR001730; Peptidase_C4.
DR   InterPro; IPR001456; Peptidase_C6.
DR   InterPro; IPR001592; Poty_coat.
DR   InterPro; IPR002540; Poty_P1.
DR   InterPro; IPR007095; RNA_pol_DS_PS.
DR   InterPro; IPR001205; RNA_pol_P3D.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF00271; helicase_C; 1.
DR   Pfam; PF00863; Peptidase_C4; 1.
DR   Pfam; PF00851; Peptidase_C6; 1.
DR   Pfam; PF00767; Poty_coat; 1.
DR   Pfam; PF01577; Poty_P1; 1.
DR   Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR   PRINTS; PR00966; NIAPOTYPTASE.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   PROSITE; PS50507; RDRP_POSITIVE; 1.
DR   PROSITE; PS50521; RDRP_VIRAL; 1.
KW   ATP-binding; Coat protein; Helicase; Hydrolase.
FT   CHAIN         1    362       P1 PROTEIN.
FT   CHAIN       363    818       HC-PRO PROTEIN.
FT   CHAIN       819   1177       P3 PROTEIN.
FT   CHAIN      1178   1229       6K1 PROTEIN.
FT   CHAIN      1230   1864       CI PROTEIN.
FT   CHAIN      1865   1917       6K2 PROTEIN.
FT   CHAIN      1918   2109       NIA-VPG PROTEIN.
FT   CHAIN      2110   2351       NIA-PRO PROTEIN.
FT   CHAIN      2352   2864       NIB PROTEIN.
FT   CHAIN      2865   3152       COAT PROTEIN.
SQ   SEQUENCE   3152 AA;  357600 MW;  EB881A928A5FA16D CRC64;

  Query Match             96.8%;  Score 30;  DB 12;  Length 3152;
  Best Local Similarity   83.3%;  Pred. No. 4.7e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              |:||||
Db        403 LVQLFH 408


RESULT 7
Q8BZC2
ID   Q8BZC2      PRELIMINARY;      PRT;  136 AA.
AC   Q8BZC2;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Aspartoacylase (Fragment).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK035910; BAC29240.1; -.
FT   NON_TER       1       1
SQ   SEQUENCE   136 AA;  15735 MW;  05B4CC681FEAB785 CRC64;

  Query Match             93.5%;  Score 29;  DB 11;  Length 136;
  Best Local Similarity   83.3%;  Pred. No. 43;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              |||:||
Db         85 LIQMFH 90


RESULT 8
Q9JIN8
ID   Q9JIN8      PRELIMINARY;      PRT;  312 AA.
AC   Q9JIN8;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Aspartoacylase ASPA.
GN   ASPA.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20299153; PubMed=10837925;
RA   Namboodiri M.A., Corigliano-Murphy A., Jiang G., Rollag M.,
RA   Provencio I.;
RT   "Murine aspartoacylase: cloning, expression and comparison with the
RT   human enzyme.";
RL   Brain Res. Mol. Brain Res. 77:285-289(2000).
DR   EMBL; AF212998; AAF76459.1; -.
DR   MGD; MGI:87914; Aspa.
DR   InterPro; IPR007036; Aste_AspA.
DR   Pfam; PF04952; Aste_AspA; 1.
SQ   SEQUENCE   312 AA;  35304 MW;  65A8C55F20F1755B CRC64;

  Query Match             93.5%;  Score 29;  DB 11;  Length 312;
  Best Local Similarity   83.3%;  Pred. No. 93;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              |||:||
Db        135 LIQMFH 140
```

```
RESULT 9
Q9DCE2
ID   Q9DCE2      PRELIMINARY;      PRT;   312 AA.
AC   Q9DCE2;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Aspartoacylase (Aminoacylase) 2.
GN   ASPA.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Kidney;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
DR   EMBL; AK002859; BAB22412.1; -.
DR   MGD; MGI:87914; Aspa.
DR   InterPro; IPR007036; Aste_AspA.
DR   Pfam; PF04952; Aste_AspA; 1.
SQ   SEQUENCE   312 AA;  35344 MW;  F09E710CC3107108 CRC64;

  Query Match             93.5%;  Score 29;  DB 11;  Length 312;
  Best Local Similarity   83.3%;  Pred. No. 93;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              |||:||
Db        135 LIQMFH 140


RESULT 10
Q8R3P0
ID   Q8R3P0      PRELIMINARY;      PRT;   312 AA.
AC   Q8R3P0;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Aspartoacylase (Aminoacylase) 2.
GN   ASPA.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC024934; AAH24934.1; -.
DR   MGD; MGI:87914; Aspa.
DR   InterPro; IPR007036; Aste_AspA.
DR   Pfam; PF04952; Aste_AspA; 1.
SQ   SEQUENCE   312 AA;  35344 MW;  7328C55F20F1755B CRC64;

  Query Match             93.5%;  Score 29;  DB 11;  Length 312;
  Best Local Similarity   83.3%;  Pred. No. 93;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              |||:||
Db        135 LIQMFH 140


RESULT 11
Q9D876
ID   Q9D876      PRELIMINARY;      PRT;   312 AA.
AC   Q9D876;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Aspartoacylase (Aminoacylase) 2.
GN   ASPA.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Small intestine;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
DR   EMBL; AK008354; BAB25624.1; -.
DR   MGD; MGI:87914; Aspa.
DR   InterPro; IPR007036; Aste_AspA.
DR   Pfam; PF04952; Aste_AspA; 1.
SQ   SEQUENCE   312 AA;  35388 MW;  0EF954DC9C1FCA08 CRC64;

  Query Match             93.5%;  Score 29;  DB 11;  Length 312;
  Best Local Similarity   83.3%;  Pred. No. 93;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              |||:||
Db        135 LIQMFH 140


RESULT 12
Q9R1T5
ID   Q9R1T5      PRELIMINARY;      PRT;   314 AA.
AC   Q9R1T5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Aspartoacylase.
```

```
GN   ASPA.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WTC; TISSUE=Brain;
RA   Kitada K., Serikawa T.;
RT   "Full length rat cDNA coding for aspartoacylase.";
RL   Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB023432; BAA82801.1; -.
DR   InterPro; IPR007036; Aste_AspA.
DR   Pfam; PF04952; Aste_AspA; 1.
SQ   SEQUENCE   314 AA;  35580 MW;  6F307333F7334F86 CRC64;

  Query Match             93.5%;  Score 29;  DB 11;  Length 314;
  Best Local Similarity   83.3%;  Pred. No. 93;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              |||:||
Db        135 LIQMFH 140


RESULT 13
Q9VX89
ID   Q9VX89      PRELIMINARY;      PRT;   379 AA.
AC   Q9VX89;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   CG12433 protein.
GN   CG12433.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003504; AAF48689.1; -.
DR   FlyBase; FBgn0030811; CG12433.
SQ   SEQUENCE   379 AA;  43974 MW;  8F24898E3F489A13 CRC64;

  Query Match             93.5%;  Score 29;  DB 5;  Length 379;
  Best Local Similarity   83.3%;  Pred. No. 1.1e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              |:||||
Db        195 LLQLFH 200


RESULT 14
Q8LQJ0
ID   Q8LQJ0      PRELIMINARY;      PRT;  1627 AA.
AC   Q8LQJ0;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   P0456E05.12 protein.
GN   P0456E05.12.
OS   Oryza sativa (japonica cultivar-group).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=39947;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Nipponbare;
RA   Sasaki T., Matsumoto T., Yamamoto K.;
RT   "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT   clone:P0456E05.";
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AP003416; BAB92613.1; -.
DR   Gramene; Q8LQJ0; -.
DR   InterPro; IPR001064; Crystallin.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR005034; DUF283.
DR   InterPro; IPR001650; Helicase_C.
DR   InterPro; IPR003100; PAZ.
DR   InterPro; IPR000999; RNase_3.
DR   InterPro; IPR007087; Znf_C2H2.
DR   Pfam; PF03368; DUF283; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   Pfam; PF02170; PAZ; 1.
DR   Pfam; PF00636; Ribonuclease_3; 1.
DR   Pfam; PF00096; zf-C2H2; 1.
DR   SMART; SM00490; HELICc; 1.
DR   SMART; SM00535; RIBOc; 2.
DR   PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; 1.
DR   PROSITE; PS50821; PAZ; 1.
DR   PROSITE; PS00517; RNASE_3_1; 1.
DR   PROSITE; PS50142; RNASE_3_2; 2.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 1.
KW   ATP-binding; Helicase; Hydrolase; Metal-binding; Zinc; Zinc-finger.
SQ   SEQUENCE   1627 AA;  180368 MW;  DF7776452C2068F0 CRC64;

  Query Match             93.5%;  Score 29;  DB 10;  Length 1627;
  Best Local Similarity   83.3%;  Pred. No. 4.4e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              |||:||
```

```
Db        226 LIQIFH 231


RESULT 15
Q9KGH7
ID   Q9KGH7      PRELIMINARY;      PRT;   179 AA.
AC   Q9KGH7;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypoxanthine-guanine phosphoribosyltransferase (EC 2.4.2.8).
GN   HPRT OR BH0084.
OS   Bacillus halodurans.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=86665;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C-125 / JCM 9153;
RX   MEDLINE=20512582; PubMed=11058132;
RA   Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA   Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA   Horikoshi K.;
RT   "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT   halodurans and genomic sequence comparison with Bacillus subtilis.";
RL   Nucleic Acids Res. 28:4317-4331(2000).
CC   -!- SIMILARITY: BELONGS TO THE PURINE/PYRIMIDINE
CC       PHOSPHORIBOSYLTRANSFERASE FAMILY.
DR   EMBL; AP001507; BAB03803.1; -.
DR   HSSP; P51900; 1HCX.
DR   InterPro; IPR005904; Hxn_phspho_trans.
DR   InterPro; IPR002375; Pr/py_rp_transf.
DR   InterPro; IPR000836; PRTransferase.
DR   Pfam; PF00156; Pribosyltran; 1.
DR   TIGRFAMs; TIGR01203; HGPRTase; 1.
DR   PROSITE; PS00103; PUR_PYR_PR_TRANSFER; 1.
KW   Glycosyltransferase; Transferase; Complete proteome.
SQ   SEQUENCE   179 AA;  20139 MW;  BA0B681C58D83BC9 CRC64;

  Query Match             90.3%;  Score 28;  DB 16;  Length 179;
  Best Local Similarity   83.3%;  Pred. No. 94;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIQLFH 6
              ||:|||
Db        110 LIELFH 115


Search completed: October 14, 2003, 14:36:50
Job time : 22.875 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 14:33:19 ; Search time 25.75 Seconds
(without alignments)
36.985 Million cell updates/sec

Title: US-09-881-490-154
Perfect score: 33
Sequence: 1 WLIQLK 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	33	100.0	6	17	AAW04065	Antifungal peptide
2	33	100.0	6	18	AAW44553	Anti-fungal peptid
3	33	100.0	6	18	AAW43743	Bactericidal/perme
4	33	100.0	6	20	AAY00530	Antifungal peptide
5	33	100.0	6	22	AAB65454	Anti-fungal peptid
6	33	100.0	7	17	AAW04025	Antifungal peptide
7	33	100.0	7	17	AAW04026	Antifungal peptide
8	33	100.0	7	18	AAW44554	Anti-fungal peptid
9	33	100.0	7	18	AAW44555	Anti-fungal peptid
10	33	100.0	7	18	AAW43746	Bactericidal/perme
11	33	100.0	7	18	AAW43744	Bactericidal/perme
12	33	100.0	7	18	AAW43745	Bactericidal/perme
13	33	100.0	7	20	AAY00531	Antifungal peptide
14	33	100.0	7	20	AAY00532	Antifungal peptide
15	33	100.0	7	22	AAB65455	Anti-fungal peptid
16	33	100.0	7	22	AAB65456	Anti-fungal peptid
17	33	100.0	8	17	AAW04027	Antifungal peptide
18	33	100.0	8	17	AAW04029	Antifungal peptide
19	33	100.0	8	18	AAW44556	Anti-fungal peptid
20	33	100.0	8	18	AAW44558	Anti-fungal peptid
21	33	100.0	8	18	AAW43748	Bactericidal/perme
22	33	100.0	8	20	AAY00535	Antifungal peptide
23	33	100.0	8	20	AAY00533	Antifungal peptide
24	33	100.0	8	22	AAB65457	Anti-fungal peptid
25	33	100.0	8	22	AAB65459	Anti-fungal peptid
26	33	100.0	9	17	AAW04030	Antifungal peptide
27	33	100.0	9	17	AAW04031	Antifungal peptide
28	33	100.0	9	18	AAW44559	Anti-fungal peptid
29	33	100.0	9	18	AAW44561	Anti-fungal peptid
30	33	100.0	9	18	AAW44563	Anti-fungal peptid
31	33	100.0	9	18	AAW44564	Anti-fungal peptid
32	33	100.0	9	18	AAW44565	Anti-fungal peptid
33	33	100.0	9	18	AAW44572	Anti-fungal peptid
34	33	100.0	9	18	AAW43754	Bactericidal/perme
35	33	100.0	9	18	AAW43753	Bactericidal/perme
36	33	100.0	9	18	AAW43749	Bactericidal/perme
37	33	100.0	9	20	AAY00536	Antifungal peptide
38	33	100.0	9	20	AAY00538	Antifungal peptide
39	33	100.0	9	20	AAY00540	Antifungal peptide
40	33	100.0	9	20	AAY00541	Antifungal peptide
41	33	100.0	9	20	AAY00542	Antifungal peptide
42	33	100.0	9	20	AAY00549	Antifungal peptide
43	33	100.0	9	21	AAY88384	Antifungal peptide
44	33	100.0	9	22	AAB65460	Anti-fungal peptid
45	33	100.0	9	22	AAB65462	Anti-fungal peptid

ALIGNMENTS

RESULT 1
AAW04065
ID AAW04065 standard; peptide; 6 AA.
XX
AC AAW04065;
XX
DT 04-NOV-1996 (first entry)
XX
DE Antifungal peptide XMP.321.
XX
KW Antifungal peptide; inhibitor; Domain III; polymorphonuclear leukocyte;
KW bactericidal/permeability-increasing protein; BPI; mammalian; PMN; fungi;
KW neutrophil; replication inhibitor; fungal infection; Aspergillus;
KW Cryptococcus; Candida; C.albicans; C.galabrat; C.krusei; C.lusitaniae;
KW C.parapsilosis; C.tropicalis; therapy.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 6
FT /note= "amidated"
XX
PN WO9608509-A1.
XX
PD 21-MAR-1996.
XX
PF 20-JUL-1995; 95WO-US09262.
XX
PR 13-JAN-1995; 95US-0372105.
PR 15-SEP-1994; 94US-0306473.
XX
PA (XOMA) XOMA CORP.

XX
PI Fadem MB, Lim E, Little RG;
XX
DR WPI; 1996-179900/18.
XX
PT Antifungal peptide(s) derived from Domain III of BPI protein - used
PT in vitro for killing or inhibiting replication of fungi, esp.
PT Candida species
XX
PS Claim 14; Page 151; 199pp; English.
XX
CC AAW04000-W04160 represent antifungal peptides. These sequences are
CC based on (or derived from) Domain III of the
CC bactericidal/permeability-increasing protein (BPI). BPI is a protein
CC that can be isolated from the granules of mammalian polymorphonuclear
CC leukocytes (PMNs or neutrophils). These antifungal peptides can be used
CC for killing, or inhibiting replication of, fungi in vitro. These
CC sequences can also be used for treatment of a fungal infection involving
CC fungi from the species Candida, Aspergillus and Cryptococcus. The
CC sequences are especially useful for treating C.albicans, C.galabrat,
CC C.krusei, C.lusitaniae, C.parapsilosis and C.tropicalis infections.
XX
SQ Sequence 6 AA;

Query Match 100.0%; Score 33; DB 17; Length 6;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WLIQLK 6
||||||
Db 1 WLIQLK 6

RESULT 2
AAW44553
ID AAW44553 standard; peptide; 6 AA.
XX
AC AAW44553;
XX
DT 27-APR-1998 (first entry)
XX
DE Anti-fungal peptide #154 based on BPI protein (residues 142-169).
XX
KW Anti-fungal peptide; bactericidal-permeability-increasing protein; BPI;
KW polymorphonuclear leukocyte; fungicide.
XX
OS Synthetic.
OS Mammalia.
XX
FH Key Location/Qualifiers
FT Modified-site 6
FT /note= "C-terminal amide"
XX
PN WO9704008-A1.
XX
PD 06-FEB-1997.
XX
PF 21-MAR-1996; 96WO-US03845.
XX
PR 20-JUL-1995; 95US-0504841.
XX
PA (XOMA) XOMA CORP.
XX
PI Fadem MB, Lim E, Little RG;
XX
DR WPI; 1997-132578/12.
XX
PT Anti-fungal peptide(s) derived from or based on domain III of
PT bactericidal-permeability-increasing protein - are used in vitro or
PT in vivo as a fungicides
XX
PS Claim 1; Page 188; 230pp; English.
XX
CC This is a specifically claimed anti-fungal peptide which is based on
CC domain III (amino acids 142-160) of bactericidal-permeability-increasing
CC protein (BPI), isolated from the granules of mammalian polymorphonuclear
CC leukocytes. It is used in compositions with diluents, carriers or
CC adjuvants to treat fungal infections in patients. It may also be used in
CC vitro to kill or inhibit the replication of fungi, such as in
CC decontaminating fluids and sterilising medical and implant devices.
XX
SQ Sequence 6 AA;

Query Match 100.0%; Score 33; DB 18; Length 6;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WLIQLK 6
||||||
Db 1 WLIQLK 6

RESULT 3
AAW43743
ID AAW43743 standard; peptide; 6 AA.
XX
AC AAW43743;
XX
DT 20-APR-1998 (first entry)
XX
DE Bactericidal/permeability increasing peptide XMP.321.
XX
KW Bactericidal/permeability increasing peptide; BPI; fusion protein;
KW bacterial infection; fungal infection; endotoxin; heparin;
KW angiogenesis; fungicidal; recombinant DNA; vector.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 6
FT /note= "Amidated"
XX
PN WO9735009-A1.
XX
PD 25-SEP-1997.
XX
PF 18-MAR-1997; 97WO-US05287.
XX
PR 22-MAR-1996; 96US-0621803.
XX
PA (XOMA) XOMA CORP.
XX
PI Better MD;
XX
DR WPI; 1997-480215/44.
XX
PT Recombinant production of bactericidal/permeability increasing
PT protein - by expression as a fusion protein in microbial host cells,
PT then cleaving the BPI peptide from the carrier
XX
PS Claim 10; Page 123; 186pp; English.
XX
CC A new recombinant DNA vector construct has been developed which encodes
CC a fusion protein and is suitable for introduction into a bacterial host.
CC The vector comprises: (a) DNA encoding at least one cationic
CC bactericidal/permeability increasing peptide (BPI), (b) DNA encoding a
CC carrier protein, and (c) DNA encoding an amino acid (aa) cleavage site
CC located between (a) and (b). The present sequence represents a
CC specifically claimed BPI peptide. The peptides have many uses including
CC the treatment of bacterial and fungal infections. BPI peptides also
CC bind to endotoxins and heparin, neutralising their effects. The
CC peptides have further been shown to inhibit angiogenesis (partly due to
CC heparin-binding activity). The fusion proteins have been found to be
CC expressed in large amounts without significant proteolysis, and in some
CC cases are actually secreted from the host cells. This allows the

```
CC   indirect production of anti-microbial BPI peptides in microbial hosts.
XX
SQ   Sequence   6 AA;

  Query Match             100.0%;  Score 33;  DB 18;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 9.3e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              ||||||
Db          1 WLIQLK 6


RESULT 4
AAY00530
ID   AAY00530 standard; Peptide; 6 AA.
XX
AC   AAY00530;
XX
DT   07-MAY-1999  (first entry)
XX
DE   Antifungal peptide XMP.321.
XX
KW   Antifungal; BPI; bactericidal/permeability increasing protein;
KW   Candida infection.
XX
OS   Synthetic.
XX
PN   US5858974-A.
XX
PD   12-JAN-1999.
XX
PF   21-MAR-1996;   96US-0621259.
XX
PR   21-MAR-1996;   96US-0621259.
PR   20-JUL-1995;   95US-0504841.
XX
PA   (XOMA ) XOMA CORP.
XX
PI   Fadem MB,  Lim E,  Little RG;
XX
DR   WPI; 1999-119956/10.
XX
PT   Antifungal peptides - comprising part of bactericidal or
PT   permeability-increasing protein sequence or related sequence
XX
PS   Disclosure; Columns 161-162; 132pp; English.
XX
CC   New peptides are provided which are based on Domain III (amino acids
CC   142-169) of human bactericidal/permeability-increasing protein (BPI).
CC   The peptides all have a C-terminal amide. More particularly, the Claims
CC   relate to: (1) a peptide that has an amino acid sequence of human BPI
CC   from position 148 to position 161 (KSKVGWLIQLFHKK) and variants of the
CC   sequence having antifungal activity; and (2) an antifungal peptide
CC   having 6-14 amino acids comprising (a) a core sequence selected from
CC   LIQL, IQLF, WLIQL, LIQLF and WLIQLF and (b) one or more cationic amino
CC   acids selected from K, R, H, Orn (ornithine) and Dab (diaminobutyric
CC   acid) at the N and/or C terminus of the core sequence. The new peptides
CC   are used for killing or inhibiting replication of fungi in vitro; and
CC   for treating fungal infections in vivo, in particular infections of
CC   Candida, Aspergillus or Cryptococcus spp., especially C. albicans, C.
CC   krusei, C. lusitaniae, C. parapsilosis or C. tropicalis. The peptide
CC   can be administered topically, intravenously, orally or as an aerosol,
CC   optionally together with a non-peptide antifungal agent.
XX
SQ   Sequence   6 AA;

  Query Match             100.0%;  Score 33;  DB 20;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 9.3e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              ||||||
Db          1 WLIQLK 6


RESULT 5
AAB65454
ID   AAB65454 standard; Peptide; 6 AA.
XX
AC   AAB65454;
XX
DT   27-MAR-2001  (first entry)
XX
DE   Anti-fungal peptide XMP.321.
XX
KW   Human; BPI; antifungal; polymorphonuclear leukocyte; neutrophil;
KW   bactericidal/permeability-increasing protein; bactericidal;
KW   fungal infection.
XX
OS   Homo sapiens.
XX
PN   US6156730-A.
XX
PD   05-DEC-2000.
XX
PF   08-JAN-1999;   99US-0227659.
XX
PR   21-MAR-1996;   96US-0621259.
PR   12-MAR-1993;   93US-0030644.
PR   15-JUL-1993;   93US-0093202.
PR   14-JAN-1994;   94US-0183222.
PR   11-MAR-1994;   94US-0209762.
PR   11-JUL-1994;   94US-0273540.
PR   15-SEP-1994;   94US-0306473.
PR   13-JAN-1995;   95US-0372105.
PR   20-JUL-1995;   95US-0504841.
XX
PA   (XOMA ) XOMA CORP.
XX
PI   Lim E,  Fadem MB,  Little RG;
XX
DR   WPI; 2001-090160/10.
XX
PT   Novel anti-fungal peptides derived from domain III of
PT   bactericidal/permeability-increasing protein useful for killing or
PT   inhibiting replication of fungi and for treating fungal infections  -
XX
PS   Example 2; Columns 165-166; 134pp; English.
XX
CC   The present invention relates to antifungal peptides (see
CC   AAB65301-B65550) derived from Domain III (amino acids 142-169) of
CC   bactericidal/permeability-increasing protein (BPI). The present sequence
CC   is one such antifungal peptide. BPI is a protein isolated from the
CC   granules of mammalian polymorphonuclear leukocytes (PMNs or
CC   neutrophils). BPI has potent bactericidal activity against a broad range
CC   of gram-negative bacteria. The peptides of the present invention are
CC   useful for killing or inhibiting replication of fungi, and treating
CC   infections caused by fungus belonging to Candida, Aspergillus,
CC   Cryptococcus species such as C.albicans, C.glabrata, C.krusei,
CC   C.lusitaniae, C.parapsilosis and C.tropicalis.
XX
SQ   Sequence   6 AA;

  Query Match             100.0%;  Score 33;  DB 22;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 9.3e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              ||||||
Db          1 WLIQLK 6


RESULT 6
AAW04025
ID   AAW04025 standard; peptide; 7 AA.
```

```
XX
AC   AAW04025;
XX
DT   31-OCT-1996  (first entry)
XX
DE   Antifungal peptide XMP.322.
XX
KW   Antifungal peptide; inhibitor; Domain III; polymorphonuclear leukocyte;
KW   bactericidal/permeability-increasing protein; BPI; mammalian; PMN; fungi;
KW   neutrophil; replication inhibitor; fungal infection; Aspergillus;
KW   Cryptococcus; Candida; C.albicans; C.galabrat; C.krusei; C.lusitaniae;
KW   C.parapsilosis; C.tropicalis; therapy.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   7
FT                   /note= "amidated"
XX
PN   WO9608509-A1.
XX
PD   21-MAR-1996.
XX
PF   20-JUL-1995;    95WO-US09262.
XX
PR   13-JAN-1995;    95US-0372105.
PR   15-SEP-1994;    94US-0306473.
XX
PA   (XOMA ) XOMA CORP.
XX
PI   Fadem MB,  Lim E,  Little RG;
XX
DR   WPI; 1996-179900/18.
XX
PT   Antifungal peptide(s) derived from Domain III of BPI protein  - used
PT   in vitro for killing or inhibiting replication of fungi, esp.
PT   Candida species
XX
PS   Claim 5; Page 152; 199pp; English.
XX
CC   AAW04000-W04160 represent antifungal peptides.  These sequences are
CC   based on (or derived from) Domain III of the
CC   bactericidal/permeability-increasing protein (BPI).  BPI is a protein
CC   that can be isolated from the granules of mammalian polymorphonuclear
CC   leukocytes (PMNs or neutrophils).  These antifungal peptides can be used
CC   for killing, or inhibiting replication of, fungi in vitro.  These
CC   sequences can also be used for treatment of a fungal infection involving
CC   fungi from the species Candida, Aspergillus and Cryptococcus.  The
CC   sequences are especially useful for treating C.albicans, C.galabrat,
CC   C.krusei, C.lusitaniae, C.parapsilosis and C.tropicalis infections.
XX
SQ   Sequence   7 AA;

  Query Match             100.0%;  Score 33;  DB 17;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 9.3e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              ||||||
Db          1 WLIQLK 6


RESULT 7
AAW04026
ID   AAW04026 standard; peptide; 7 AA.
XX
AC   AAW04026;
XX
DT   31-OCT-1996  (first entry)
XX
DE   Antifungal peptide XMP.323.
XX
KW   Antifungal peptide; inhibitor; Domain III; polymorphonuclear leukocyte;
KW   bactericidal/permeability-increasing protein; BPI; mammalian; PMN; fungi;
KW   neutrophil; replication inhibitor; fungal infection; Aspergillus;
KW   Cryptococcus; Candida; C.albicans; C.galabrat; C.krusei; C.lusitaniae;
KW   C.parapsilosis; C.tropicalis; therapy.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   7
FT                   /note= "amidated"
XX
PN   WO9608509-A1.
XX
PD   21-MAR-1996.
XX
PF   20-JUL-1995;    95WO-US09262.
XX
PR   13-JAN-1995;    95US-0372105.
PR   15-SEP-1994;    94US-0306473.
XX
PA   (XOMA ) XOMA CORP.
XX
PI   Fadem MB,  Lim E,  Little RG;
XX
DR   WPI; 1996-179900/18.
XX
PT   Antifungal peptide(s) derived from Domain III of BPI protein  - used
PT   in vitro for killing or inhibiting replication of fungi, esp.
PT   Candida species
XX
PS   Claim 5; Page 152; 199pp; English.
XX
CC   AAW04000-W04160 represent antifungal peptides.  These sequences are
CC   based on (or derived from) Domain III of the
CC   bactericidal/permeability-increasing protein (BPI).  BPI is a protein
CC   that can be isolated from the granules of mammalian polymorphonuclear
CC   leukocytes (PMNs or neutrophils).  These antifungal peptides can be used
CC   for killing, or inhibiting replication of, fungi in vitro.  These
CC   sequences can also be used for treatment of a fungal infection involving
CC   fungi from the species Candida, Aspergillus and Cryptococcus.  The
CC   sequences are especially useful for treating C.albicans, C.galabrat,
CC   C.krusei, C.lusitaniae, C.parapsilosis and C.tropicalis infections.
XX
SQ   Sequence   7 AA;

  Query Match             100.0%;  Score 33;  DB 17;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 9.3e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              ||||||
Db          2 WLIQLK 7


RESULT 8
AAW44554
ID   AAW44554 standard; peptide; 7 AA.
XX
AC   AAW44554;
XX
DT   27-APR-1998  (first entry)
XX
DE   Anti-fungal peptide #155 based on BPI protein (residues 142-169).
XX
KW   Anti-fungal peptide; bactericidal-permeability-increasing protein; BPI;
KW   polymorphonuclear leukocyte; fungicide.
XX
OS   Synthetic.
OS   Mammalia.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   7
FT                   /note= "C-terminal amide"
```

XX
PN WO9704008-A1.
XX
PD 06-FEB-1997.
XX
PF 21-MAR-1996; 96WO-US03845.
XX
PR 20-JUL-1995; 95US-0504841.
XX
PA (XOMA) XOMA CORP.
XX
PI Fadem MB, Lim E, Little RG;
XX
DR WPI; 1997-132578/12.
XX
PT Anti-fungal peptide(s) derived from or based on domain III of
PT bactericidal-permeability-increasing protein - are used in vitro or
PT in vivo as a fungicides
XX
PS Claim 1; Page 189; 230pp; English.
XX
CC This is a specifically claimed anti-fungal peptide which is based on
CC domain III (amino acids 142-160) of bactericidal-permeability-increasing
CC protein (BPI), isolated from the granules of mammalian polymorphonuclear
CC leukocytes. It is used in compositions with diluents, carriers or
CC adjuvants to treat fungal infections in patients. It may also be used in
CC vitro to kill or inhibit the replication of fungi, such as in
CC decontaminating fluids and sterilising medical and implant devices.
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 33; DB 18; Length 7;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WLIQLK 6
||||||
Db 1 WLIQLK 6

RESULT 9
AAW44555
ID AAW44555 standard; peptide; 7 AA.
XX
AC AAW44555;
XX
DT 27-APR-1998 (first entry)
XX
DE Anti-fungal peptide #156 based on BPI protein (residues 142-169).
XX
KW Anti-fungal peptide; bactericidal-permeability-increasing protein; BPI;
KW polymorphonuclear leukocyte; fungicide.
XX
OS Synthetic.
OS Mammalia.
XX
FH Key Location/Qualifiers
FT Modified-site 7
FT /note= "C-terminal amide"
XX
PN WO9704008-A1.
XX
PD 06-FEB-1997.
XX
PF 21-MAR-1996; 96WO-US03845.
XX
PR 20-JUL-1995; 95US-0504841.
XX
PA (XOMA) XOMA CORP.
XX
PI Fadem MB, Lim E, Little RG;
XX
DR WPI; 1997-132578/12.
XX
PT Anti-fungal peptide(s) derived from or based on domain III of
PT bactericidal-permeability-increasing protein - are used in vitro or
PT in vivo as a fungicides
XX
PS Claim 1; Page 189; 230pp; English.
XX
CC This is a specifically claimed anti-fungal peptide which is based on
CC domain III (amino acids 142-160) of bactericidal-permeability-increasing
CC protein (BPI), isolated from the granules of mammalian polymorphonuclear
CC leukocytes. It is used in compositions with diluents, carriers or
CC adjuvants to treat fungal infections in patients. It may also be used in
CC vitro to kill or inhibit the replication of fungi, such as in
CC decontaminating fluids and sterilising medical and implant devices.
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 33; DB 18; Length 7;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WLIQLK 6
||||||
Db 2 WLIQLK 7

RESULT 10
AAW43746
ID AAW43746 standard; peptide; 7 AA.
XX
AC AAW43746;
XX
DT 20-APR-1998 (first entry)
XX
DE Bactericidal/permeability increasing peptide XMP.324.
XX
KW Bactericidal/permeability increasing peptide; BPI; fusion protein;
KW bacterial infection; fungal infection; endotoxin; heparin;
KW angiogenesis; fungicidal; recombinant DNA; vector.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 7
FT /note= "Amidated"
XX
PN WO9735009-A1.
XX
PD 25-SEP-1997.
XX
PF 18-MAR-1997; 97WO-US05287.
XX
PR 22-MAR-1996; 96US-0621803.
XX
PA (XOMA) XOMA CORP.
XX
PI Better MD;
XX
DR WPI; 1997-480215/44.
XX
PT Recombinant production of bactericidal/permeability increasing
PT protein - by expression as a fusion protein in microbial host cells,
PT then cleaving the BPI peptide from the carrier
XX
PS Claim 10; Page 124; 186pp; English.
XX
CC A new recombinant DNA vector construct has been developed which encodes
CC a fusion protein and is suitable for introduction into a bacterial host.
CC The vector comprises: (a) DNA encoding at least one cationic
CC bactericidal/permeability increasing peptide (BPI), (b) DNA encoding a
CC carrier protein, and (c) DNA encoding an amino acid (aa) cleavage site
CC located between (a) and (b). The present sequence represents a

CC specifically claimed BPI peptide. The peptides have many uses including
CC the treatment of bacterial and fungal infections. BPI peptides also
CC bind to endotoxins and heparin, neutralising their effects. The
CC peptides have further been shown to inhibit angiogenesis (partly due to
CC heparin-binding activity). The fusion proteins have been found to be
CC expressed in large amounts without significant proteolysis, and in some
CC cases are actually secreted from the host cells. This allows the
CC indirect production of anti-microbial BPI peptides in microbial hosts.
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 33; DB 18; Length 7;
Best Local Similarity 100.0%; Pred. No. 9.3e-05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WLIQLK 6
||||||
Db 2 WLIQLK 7

RESULT 11
AAW43744
ID AAW43744 standard; peptide; 7 AA.
XX
AC AAW43744;
XX
DT 20-APR-1998 (first entry)
XX
DE Bactericidal/permeability increasing peptide XMP.322.
XX
KW Bactericidal/permeability increasing peptide; BPI; fusion protein;
KW bacterial infection; fungal infection; endotoxin; heparin;
KW angiogenesis; fungicidal; recombinant DNA; vector.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 7
FT /note= "Amidated"
XX
PN WO9735009-A1.
XX
PD 25-SEP-1997.
XX
PF 18-MAR-1997; 97WO-US05287.
XX
PR 22-MAR-1996; 96US-0621803.
XX
PA (XOMA) XOMA CORP.
XX
PI Better MD;
XX
DR WPI; 1997-480215/44.
XX
PT Recombinant production of bactericidal/permeability increasing
PT protein - by expression as a fusion protein in microbial host cells,
PT then cleaving the BPI peptide from the carrier
XX
PS Claim 10; Page 124; 186pp; English.
XX
CC A new recombinant DNA vector construct has been developed which encodes
CC a fusion protein and is suitable for introduction into a bacterial host.
CC The vector comprises: (a) DNA encoding at least one cationic
CC bactericidal/permeability increasing peptide (BPI), (b) DNA encoding a
CC carrier protein, and (c) DNA encoding an amino acid (aa) cleavage site
CC located between (a) and (b). The present sequence represents a
CC specifically claimed BPI peptide. The peptides have many uses including
CC the treatment of bacterial and fungal infections. BPI peptides also
CC bind to endotoxins and heparin, neutralising their effects. The
CC peptides have further been shown to inhibit angiogenesis (partly due to
CC heparin-binding activity). The fusion proteins have been found to be
CC expressed in large amounts without significant proteolysis, and in some
CC cases are actually secreted from the host cells. This allows the
CC indirect production of anti-microbial BPI peptides in microbial hosts.
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 33; DB 18; Length 7;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WLIQLK 6
||||||
Db 1 WLIQLK 6

RESULT 12
AAW43745
ID AAW43745 standard; peptide; 7 AA.
XX
AC AAW43745;
XX
DT 20-APR-1998 (first entry)
XX
DE Bactericidal/permeability increasing peptide XMP.323.
XX
KW Bactericidal/permeability increasing peptide; BPI; fusion protein;
KW bacterial infection; fungal infection; endotoxin; heparin;
KW angiogenesis; fungicidal; recombinant DNA; vector.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 7
FT /note= "Amidated"
XX
PN WO9735009-A1.
XX
PD 25-SEP-1997.
XX
PF 18-MAR-1997; 97WO-US05287.
XX
PR 22-MAR-1996; 96US-0621803.
XX
PA (XOMA) XOMA CORP.
XX
PI Better MD;
XX
DR WPI; 1997-480215/44.
XX
PT Recombinant production of bactericidal/permeability increasing
PT protein - by expression as a fusion protein in microbial host cells,
PT then cleaving the BPI peptide from the carrier
XX
PS Claim 10; Page 124; 186pp; English.
XX
CC A new recombinant DNA vector construct has been developed which encodes
CC a fusion protein and is suitable for introduction into a bacterial host.
CC The vector comprises: (a) DNA encoding at least one cationic
CC bactericidal/permeability increasing peptide (BPI), (b) DNA encoding a
CC carrier protein, and (c) DNA encoding an amino acid (aa) cleavage site
CC located between (a) and (b). The present sequence represents a
CC specifically claimed BPI peptide. The peptides have many uses including
CC the treatment of bacterial and fungal infections. BPI peptides also
CC bind to endotoxins and heparin, neutralising their effects. The
CC peptides have further been shown to inhibit angiogenesis (partly due to
CC heparin-binding activity). The fusion proteins have been found to be
CC expressed in large amounts without significant proteolysis, and in some
CC cases are actually secreted from the host cells. This allows the
CC indirect production of anti-microbial BPI peptides in microbial hosts.
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 33; DB 18; Length 7;

```
  Best Local Similarity   100.0%;  Pred. No. 9.3e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              ||||:|
Db          2 WLIQLK 7


RESULT 13
AAY00531
ID   AAY00531 standard; Peptide; 7 AA.
XX
AC   AAY00531;
XX
DT   07-MAY-1999  (first entry)
XX
DE   Antifungal peptide XMP.322.
XX
KW   Antifungal; BPI; bactericidal/permeability increasing protein;
KW   Candida infection.
XX
OS   Synthetic.
XX
PN   US5858974-A.
XX
PD   12-JAN-1999.
XX
PF   21-MAR-1996;   96US-0621259.
XX
PR   21-MAR-1996;   96US-0621259.
PR   20-JUL-1995;   95US-0504841.
XX
PA   (XOMA ) XOMA CORP.
XX
PI   Fadem MB,  Lim E,  Little RG;
XX
DR   WPI; 1999-119956/10.
XX
PT   Antifungal peptides - comprising part of bactericidal or
PT   permeability-increasing protein sequence or related sequence
XX
PS   Disclosure; Columns 161-162; 132pp; English.
XX
CC   New peptides are provided which are based on Domain III (amino acids
CC   142-169) of human bactericidal/permeability-increasing protein (BPI).
CC   The peptides all have a C-terminal amide. More particularly, the Claims
CC   relate to: (1) a peptide that has an amino acid sequence of human BPI
CC   from position 148 to position 161 (KSKVGWLIQLFHKK) and variants of the
CC   sequence having antifungal activity; and (2) an antifungal peptide
CC   having 6-14 amino acids comprising (a) a core sequence selected from
CC   LIQL, IQLF, WLIQL, LIQLF and WLIQLF and (b) one or more cationic amino
CC   acids selected from K, R, H, Orn (ornithine) and Dab (diaminobutyric
CC   acid) at the N and/or C terminus of the core sequence. The new peptides
CC   are used for killing or inhibiting replication of fungi in vitro; and
CC   for treating fungal infections in vivo, in particular infections of
CC   Candida, Aspergillus or Cryptococcus spp., especially C. albicans, C.
CC   krusei, C. lusitaniae, C. parapsilosis or C. tropicalis. The peptide
CC   can be administered topically, intravenously, orally or as an aerosol,
CC   optionally together with a non-peptide antifungal agent.
XX
SQ   Sequence   7 AA;

  Query Match             100.0%;  Score 33;  DB 20;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 9.3e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              ||||||
Db          1 WLIQLK 6


RESULT 14
AAY00532
ID   AAY00532 standard; Peptide; 7 AA.
XX
AC   AAY00532;
XX
DT   07-MAY-1999  (first entry)
XX
DE   Antifungal peptide XMP.323.
XX
KW   Antifungal; BPI; bactericidal/permeability increasing protein;
KW   Candida infection.
XX
OS   Synthetic.
XX
PN   US5858974-A.
XX
PD   12-JAN-1999.
XX
PF   21-MAR-1996;   96US-0621259.
XX
PR   21-MAR-1996;   96US-0621259.
PR   20-JUL-1995;   95US-0504841.
XX
PA   (XOMA ) XOMA CORP.
XX
PI   Fadem MB,  Lim E,  Little RG;
XX
DR   WPI; 1999-119956/10.
XX
PT   Antifungal peptides - comprising part of bactericidal or
PT   permeability-increasing protein sequence or related sequence
XX
PS   Disclosure; Columns 161-162; 132pp; English.
XX
CC   New peptides are provided which are based on Domain III (amino acids
CC   142-169) of human bactericidal/permeability-increasing protein (BPI).
CC   The peptides all have a C-terminal amide. More particularly, the Claims
CC   relate to: (1) a peptide that has an amino acid sequence of human BPI
CC   from position 148 to position 161 (KSKVGWLIQLFHKK) and variants of the
CC   sequence having antifungal activity; and (2) an antifungal peptide
CC   having 6-14 amino acids comprising (a) a core sequence selected from
CC   LIQL, IQLF, WLIQL, LIQLF and WLIQLF and (b) one or more cationic amino
CC   acids selected from K, R, H, Orn (ornithine) and Dab (diaminobutyric
CC   acid) at the N and/or C terminus of the core sequence. The new peptides
CC   are used for killing or inhibiting replication of fungi in vitro; and
CC   for treating fungal infections in vivo, in particular infections of
CC   Candida, Aspergillus or Cryptococcus spp., especially C. albicans, C.
CC   krusei, C. lusitaniae, C. parapsilosis or C. tropicalis. The peptide
CC   can be administered topically, intravenously, orally or as an aerosol,
CC   optionally together with a non-peptide antifungal agent.
XX
SQ   Sequence   7 AA;

  Query Match             100.0%;  Score 33;  DB 20;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 9.3e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              ||||||
Db          2 WLIQLK 7


RESULT 15
AAB65455
ID   AAB65455 standard; Peptide; 7 AA.
XX
AC   AAB65455;
XX
DT   27-MAR-2001  (first entry)
XX
DE   Anti-fungal peptide XMP.322.
XX
KW   Human; BPI; antifungal; polymorphonuclear leukocyte; neutrophil;
KW   bactericidal/permeability-increasing protein; bactericidal;
```

KW fungal infection.
XX
OS Homo sapiens.
XX
PN US6156730-A.
XX
PD 05-DEC-2000.
XX
PF 08-JAN-1999; 99US-0227659.
XX
PR 21-MAR-1996; 96US-0621259.
PR 12-MAR-1993; 93US-0030644.
PR 15-JUL-1993; 93US-0093202.
PR 14-JAN-1994; 94US-0183222.
PR 11-MAR-1994; 94US-0209762.
PR 11-JUL-1994; 94US-0273540.
PR 15-SEP-1994; 94US-0306473.
PR 13-JAN-1995; 95US-0372105.
PR 20-JUL-1995; 95US-0504841.
XX
PA (XOMA) XOMA CORP.
XX
PI Lim E, Fadem MB, Little RG;
XX
DR WPI; 2001-090160/10.
XX
PT Novel anti-fungal peptides derived from domain III of
PT bactericidal/permeability-increasing protein useful for killing or
PT inhibiting replication of fungi and for treating fungal infections -
XX
PS Example 2; Columns 165-166; 134pp; English.
XX
CC The present invention relates to antifungal peptides (see
CC AAB65301-B65550) derived from Domain III (amino acids 142-169) of
CC bactericidal/permeability-increasing protein (BPI). The present sequence
CC is one such antifungal peptide. BPI is a protein isolated from the
CC granules of mammalian polymorphonuclear leukocytes (PMNs or
CC neutrophils). BPI has potent bactericidal activity against a broad range
CC of gram-negative bacteria. The peptides of the present invention are
CC useful for killing or inhibiting replication of fungi, and treating
CC infections caused by fungus belonging to Candida, Aspergillus,
CC Cryptococcus species such as C.albicans, C.glabrata, C.krusei,
CC C.lusitaniae, C.parapsilosis and C.tropicalis.
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 33; DB 22; Length 7;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WLIQLK 6
||||||
Db 1 WLIQLK 6

Search completed: October 14, 2003, 14:41:31
Job time : 26.9167 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 14:33:23 ; Search time 8.5 Seconds
(without alignments)
29.866 Million cell updates/sec

Title: US-09-881-490-154
Perfect score: 33
Sequence: 1 WLIQLK 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	33	100.0	6	2	US-08-621-803-187	Sequence 187, App
2	33	100.0	6	2	US-08-621-259A-154	Sequence 154, App
3	33	100.0	6	3	US-09-217-352-187	Sequence 187, App
4	33	100.0	6	5	PCT-US95-09262-154	Sequence 154, App
5	33	100.0	7	2	US-08-621-803-188	Sequence 188, App
6	33	100.0	7	2	US-08-621-803-189	Sequence 189, App
7	33	100.0	7	2	US-08-621-803-190	Sequence 190, App
8	33	100.0	7	2	US-08-621-259A-155	Sequence 155, App
9	33	100.0	7	2	US-08-621-259A-156	Sequence 156, App
10	33	100.0	7	3	US-09-217-352-188	Sequence 188, App
11	33	100.0	7	3	US-09-217-352-189	Sequence 189, App
12	33	100.0	7	3	US-09-217-352-190	Sequence 190, App
13	33	100.0	7	5	PCT-US95-09262-155	Sequence 155, App
14	33	100.0	7	5	PCT-US95-09262-156	Sequence 156, App
15	33	100.0	8	2	US-08-621-803-192	Sequence 192, App
16	33	100.0	8	2	US-08-621-259A-157	Sequence 157, App
17	33	100.0	8	2	US-08-621-259A-159	Sequence 159, App
18	33	100.0	8	3	US-09-217-352-192	Sequence 192, App
19	33	100.0	8	5	PCT-US95-09262-157	Sequence 157, App
20	33	100.0	8	5	PCT-US95-09262-159	Sequence 159, App
21	33	100.0	9	2	US-08-621-803-193	Sequence 193, App
22	33	100.0	9	2	US-08-621-803-197	Sequence 197, App
23	33	100.0	9	2	US-08-621-803-198	Sequence 198, App
24	33	100.0	9	2	US-08-621-259A-160	Sequence 160, App
25	33	100.0	9	2	US-08-621-259A-162	Sequence 162, App
26	33	100.0	9	2	US-08-621-259A-164	Sequence 164, App
27	33	100.0	9	2	US-08-621-259A-165	Sequence 165, App
28	33	100.0	9	2	US-08-621-259A-166	Sequence 166, App
29	33	100.0	9	2	US-08-621-259A-173	Sequence 173, App
30	33	100.0	9	3	US-09-217-352-193	Sequence 193, App
31	33	100.0	9	3	US-09-217-352-197	Sequence 197, App
32	33	100.0	9	3	US-09-217-352-198	Sequence 198, App
33	33	100.0	9	5	PCT-US95-09262-160	Sequence 160, App
34	33	100.0	9	5	PCT-US95-09262-162	Sequence 162, App
35	33	100.0	9	5	PCT-US95-09262-164	Sequence 164, App
36	33	100.0	9	5	PCT-US95-09262-165	Sequence 165, App
37	33	100.0	9	5	PCT-US95-09262-166	Sequence 166, App
38	33	100.0	9	5	PCT-US95-09262-173	Sequence 173, App
39	33	100.0	10	2	US-08-621-803-204	Sequence 204, App
40	33	100.0	10	2	US-08-621-803-205	Sequence 205, App
41	33	100.0	10	2	US-08-621-259A-179	Sequence 179, App
42	33	100.0	10	2	US-08-621-259A-180	Sequence 180, App
43	33	100.0	10	3	US-09-217-352-204	Sequence 204, App
44	33	100.0	10	3	US-09-217-352-205	Sequence 205, App
45	33	100.0	10	5	PCT-US95-09262-179	Sequence 179, App

ALIGNMENTS

```
RESULT 1
US-08-621-803-187
; Sequence 187, Application US/08621803
; Patent No. 5851802
;  GENERAL INFORMATION:
;    APPLICANT:  Better, Marc D.
;    TITLE OF INVENTION:  Methods for Recombinant Microbial Production of
;    TITLE OF INVENTION:  Fusion Proteins and BPI-Derived Peptides
;    NUMBER OF SEQUENCES:  265
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,803
;      FILING DATE:  22-MAR-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25,447
;      REFERENCE/DOCKET NUMBER:  27129/33199
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  187:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.321"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-08-621-803-187

  Query Match             100.0%;  Score 33;  DB 2;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 2.5e+05;
```

```
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              ||||||
Db          1 WLIQLK 6


RESULT 2
US-08-621-259A-154
; Sequence 154, Application US/08621259A
; Patent No. 5858974
;  GENERAL INFORMATION:
;    APPLICANT:  Little II, Roger G
;    APPLICANT:  Lim, Edward
;    APPLICANT:  Fadem, Mitchell B.
;    TITLE OF INVENTION:  Anti-Fungal Peptides
;    NUMBER OF SEQUENCES:  252
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McAndrews, Held & Malloy, Ltd.
;      STREET:  500 West Madison Street
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60661
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,259A
;      FILING DATE:  21-MAR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/504,841
;      FILING DATE:  20-JUL-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McNicholas, Janet M.
;      REGISTRATION NUMBER:  32,918
;      REFERENCE/DOCKET NUMBER:  11021US02
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/707-8889
;      TELEFAX:  312/707-9155
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  154:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.321"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-08-621-259A-154

  Query Match             100.0%;  Score 33;  DB 2;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              ||||||
Db          1 WLIQLK 6


RESULT 3
US-09-217-352-187
; Sequence 187, Application US/09217352
; Patent No. 6274344
;  GENERAL INFORMATION:
;    APPLICANT:  Better, Marc D.
;    TITLE OF INVENTION:  Methods for Recombinant Microbial Production of
;    TITLE OF INVENTION:  Fusion Proteins and BPI-Derived Peptides
;    NUMBER OF SEQUENCES:  265
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/217,352
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/621,803
;      FILING DATE:  22-MAR-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25,447
;      REFERENCE/DOCKET NUMBER:  27129/33199
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  187:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.321"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-09-217-352-187

  Query Match             100.0%;  Score 33;  DB 3;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 2.5e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              ||||||
Db          1 WLIQLK 6


RESULT 4
PCT-US95-09262-154
; Sequence 154, Application PC/TUS9509262
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  Anti-Fungal Peptides
;    NUMBER OF SEQUENCES:  206
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
```

```
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US95/09262
;       FILING DATE:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/372,105
;       FILING DATE:  13-JAN-95
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/306,473
;       FILING DATE:  15-SEP-94
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/273,540
;       FILING DATE:  11-JUL-94
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/209,762
;       FILING DATE:  11-MAR-94
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/183,222
;       FILING DATE:  14-JAN-94
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/093,202
;       FILING DATE:  15-JUL-93
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/030,644
;       FILING DATE:  12-MAR-93
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Borun, Michael F.
;       REGISTRATION NUMBER:  25,447
;       REFERENCE/DOCKET NUMBER:  27129/10040
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  312/474-6300
;       TELEFAX:  312/474-0448
;       TELEX:  25-3856
;   INFORMATION FOR SEQ ID NO:  154:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  6 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
;     FEATURE:
;       NAME/KEY:  misc_feature
;       OTHER INFORMATION:  "XMP.321"
;     FEATURE:
;       NAME/KEY:  Modified-site
;       LOCATION:  C-Terminus
;       OTHER INFORMATION:  /label= Amidation
;       OTHER INFORMATION:  /note= "The C-Terminus is Amidated"
PCT-US95-09262-154

  Query Match             100.0%;  Score 33;  DB 5;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 2.5e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              ||||||
Db          1 WLIQLK 6


RESULT 5
US-08-621-803-188
; Sequence 188, Application US/08621803
; Patent No. 5851802
;  GENERAL INFORMATION:
;    APPLICANT:  Better, Marc D.
;    TITLE OF INVENTION:  Methods for Recombinant Microbial Production of
;    TITLE OF INVENTION:  Fusion Proteins and BPI-Derived Peptides
;    NUMBER OF SEQUENCES:  265
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,803
;      FILING DATE:  22-MAR-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25,447
;      REFERENCE/DOCKET NUMBER:  27129/33199
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  188:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.322"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-08-621-803-188

  Query Match             100.0%;  Score 33;  DB 2;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 2.5e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              ||||||
Db          1 WLIQLK 6


RESULT 6
US-08-621-803-189
; Sequence 189, Application US/08621803
; Patent No. 5851802
;  GENERAL INFORMATION:
;    APPLICANT:  Better, Marc D.
;    TITLE OF INVENTION:  Methods for Recombinant Microbial Production of
;    TITLE OF INVENTION:  Fusion Proteins and BPI-Derived Peptides
;    NUMBER OF SEQUENCES:  265
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,803
;      FILING DATE:  22-MAR-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25,447
;      REFERENCE/DOCKET NUMBER:  27129/33199
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
```

```
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  189:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.323"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-08-621-803-189

  Query Match             100.0%;  Score 33;  DB 2;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 2.5e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              ||||||
Db          2 WLIQLK 7


RESULT 7
US-08-621-803-190
; Sequence 190, Application US/08621803
; Patent No. 5851802
;  GENERAL INFORMATION:
;    APPLICANT:  Better, Marc D.
;    TITLE OF INVENTION:  Methods for Recombinant Microbial Production of
;    TITLE OF INVENTION:  Fusion Proteins and BPI-Derived Peptides
;    NUMBER OF SEQUENCES:  265
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,803
;      FILING DATE:  22-MAR-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25,447
;      REFERENCE/DOCKET NUMBER:  27129/33199
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  190:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.324"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-08-621-803-190

  Query Match             100.0%;  Score 33;  DB 2;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              ||||||
Db          2 WLIQLK 7


RESULT 8
US-08-621-259A-155
; Sequence 155, Application US/08621259A
; Patent No. 5858974
;  GENERAL INFORMATION:
;    APPLICANT:  Little II, Roger G
;    APPLICANT:  Lim, Edward
;    APPLICANT:  Fadem, Mitchell B.
;    TITLE OF INVENTION:  Anti-Fungal Peptides
;    NUMBER OF SEQUENCES:  252
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McAndrews, Held & Malloy, Ltd.
;      STREET:  500 West Madison Street
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60661
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,259A
;      FILING DATE:  21-MAR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/504,841
;      FILING DATE:  20-JUL-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McNicholas, Janet M.
;      REGISTRATION NUMBER:  32,918
;      REFERENCE/DOCKET NUMBER:  11021US02
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/707-8889
;      TELEFAX:  312/707-9155
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  155:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.322"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-08-621-259A-155

  Query Match             100.0%;  Score 33;  DB 2;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 2.5e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              ||||||
Db          1 WLIQLK 6


RESULT 9
US-08-621-259A-156
```

```
; Sequence 156, Application US/08621259A
; Patent No. 5858974
;  GENERAL INFORMATION:
;    APPLICANT:  Little II, Roger G
;    APPLICANT:  Lim, Edward
;    APPLICANT:  Fadem, Mitchell B.
;    TITLE OF INVENTION:  Anti-Fungal Peptides
;    NUMBER OF SEQUENCES:  252
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McAndrews, Held & Malloy, Ltd.
;      STREET:  500 West Madison Street
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60661
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,259A
;      FILING DATE:  21-MAR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/504,841
;      FILING DATE:  20-JUL-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McNicholas, Janet M.
;      REGISTRATION NUMBER:  32,918
;      REFERENCE/DOCKET NUMBER:  11021US02
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/707-8889
;      TELEFAX:  312/707-9155
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  156:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.323"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-08-621-259A-156

  Query Match             100.0%;  Score 33;  DB 2;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              ||||||
Db          2 WLIQLK 7


RESULT 10
US-09-217-352-188
; Sequence 188, Application US/09217352
; Patent No. 6274344
;  GENERAL INFORMATION:
;    APPLICANT:  Better, Marc D.
;    TITLE OF INVENTION:  Methods for Recombinant Microbial Production of
;    TITLE OF INVENTION:  Fusion Proteins and BPI-Derived Peptides
;    NUMBER OF SEQUENCES:  265
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/217,352
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/621,803
;      FILING DATE:  22-MAR-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25,447
;      REFERENCE/DOCKET NUMBER:  27129/33199
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  188:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.322"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-09-217-352-188

  Query Match             100.0%;  Score 33;  DB 3;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              ||||||
Db          1 WLIQLK 6


RESULT 11
US-09-217-352-189
; Sequence 189, Application US/09217352
; Patent No. 6274344
;  GENERAL INFORMATION:
;    APPLICANT:  Better, Marc D.
;    TITLE OF INVENTION:  Methods for Recombinant Microbial Production of
;    TITLE OF INVENTION:  Fusion Proteins and BPI-Derived Peptides
;    NUMBER OF SEQUENCES:  265
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/217,352
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/621,803
;      FILING DATE:  22-MAR-1996
```

```
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25,447
;      REFERENCE/DOCKET NUMBER:  27129/33199
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  189:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.323"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-09-217-352-189

  Query Match             100.0%;  Score 33;  DB 3;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              ||||||
Db          2 WLIQLK 7


RESULT 12
US-09-217-352-190
; Sequence 190, Application US/09217352
; Patent No. 6274344
;  GENERAL INFORMATION:
;    APPLICANT:  Better, Marc D.
;    TITLE OF INVENTION:  Methods for Recombinant Microbial Production of
;    TITLE OF INVENTION:  Fusion Proteins and BPI-Derived Peptides
;    NUMBER OF SEQUENCES:  265
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/217,352
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/621,803
;      FILING DATE:  22-MAR-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25,447
;      REFERENCE/DOCKET NUMBER:  27129/33199
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  190:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.324"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-09-217-352-190

  Query Match             100.0%;  Score 33;  DB 3;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              ||||||
Db          2 WLIQLK 7


RESULT 13
PCT-US95-09262-155
; Sequence 155, Application PC/TUS9509262
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  Anti-Fungal Peptides
;    NUMBER OF SEQUENCES:  206
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/09262
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/372,105
;      FILING DATE:  13-JAN-95
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/306,473
;      FILING DATE:  15-SEP-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/273,540
;      FILING DATE:  11-JUL-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/209,762
;      FILING DATE:  11-MAR-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/183,222
;      FILING DATE:  14-JAN-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/093,202
;      FILING DATE:  15-JUL-93
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/030,644
;      FILING DATE:  12-MAR-93
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25,447
;      REFERENCE/DOCKET NUMBER:  27129/10040
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  155:
;    SEQUENCE CHARACTERISTICS:
```

```
;      LENGTH:  7 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.322"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated"
PCT-US95-09262-155

  Query Match             100.0%;  Score 33;  DB 5;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 2.5e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              ||||||
Db          1 WLIQLK 6


RESULT 14
PCT-US95-09262-156
; Sequence 156, Application PC/TUS9509262
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  Anti-Fungal Peptides
;    NUMBER OF SEQUENCES:  206
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/09262
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/372,105
;      FILING DATE:  13-JAN-95
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/306,473
;      FILING DATE:  15-SEP-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/273,540
;      FILING DATE:  11-JUL-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/209,762
;      FILING DATE:  11-MAR-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/183,222
;      FILING DATE:  14-JAN-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/093,202
;      FILING DATE:  15-JUL-93
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/030,644
;      FILING DATE:  12-MAR-93
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25,447
;      REFERENCE/DOCKET NUMBER:  27129/10040
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  156:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.323"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated"
PCT-US95-09262-156

  Query Match             100.0%;  Score 33;  DB 5;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 2.5e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              ||||||
Db          2 WLIQLK 7


RESULT 15
US-08-621-803-192
; Sequence 192, Application US/08621803
; Patent No. 5851802
;  GENERAL INFORMATION:
;    APPLICANT:  Better, Marc D.
;    TITLE OF INVENTION:  Methods for Recombinant Microbial Production of
;    TITLE OF INVENTION:  Fusion Proteins and BPI-Derived Peptides
;    NUMBER OF SEQUENCES:  265
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,803
;      FILING DATE:  22-MAR-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25,447
;      REFERENCE/DOCKET NUMBER:  27129/33199
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  192:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  8 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.326"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-08-621-803-192
```

Query Match 100.0%; Score 33; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 WLIQLK 6
              ||||||
Db          3 WLIQLK 8
```

Search completed: October 14, 2003, 14:42:45
Job time : 8.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 14:34:14 ; Search time 14.625 Seconds
(without alignments)
66.104 Million cell updates/sec

Title: US-09-881-490-154
Perfect score: 33
Sequence: 1 WLIQLK 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 600653 seqs, 161128416 residues

Total number of hits satisfying chosen parameters: 600653

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	33	100.0	6	9	US-09-765-527-187	Sequence 187, App
2	33	100.0	6	9	US-09-881-490-154	Sequence 154, App
3	33	100.0	7	9	US-09-765-527-188	Sequence 188, App
4	33	100.0	7	9	US-09-765-527-189	Sequence 189, App
5	33	100.0	7	9	US-09-765-527-190	Sequence 190, App
6	33	100.0	7	9	US-09-881-490-155	Sequence 155, App
7	33	100.0	7	9	US-09-881-490-156	Sequence 156, App
8	33	100.0	8	9	US-09-765-527-192	Sequence 192, App
9	33	100.0	8	9	US-09-881-490-157	Sequence 157, App
10	33	100.0	8	9	US-09-881-490-159	Sequence 159, App
11	33	100.0	9	9	US-09-765-527-193	Sequence 193, App
12	33	100.0	9	9	US-09-765-527-197	Sequence 197, App
13	33	100.0	9	9	US-09-765-527-198	Sequence 198, App
14	33	100.0	9	9	US-09-881-490-160	Sequence 160, App
15	33	100.0	9	9	US-09-881-490-162	Sequence 162, App
16	33	100.0	9	9	US-09-881-490-164	Sequence 164, App
17	33	100.0	9	9	US-09-881-490-165	Sequence 165, App
18	33	100.0	9	9	US-09-881-490-166	Sequence 166, App
19	33	100.0	9	9	US-09-881-490-173	Sequence 173, App
20	33	100.0	10	9	US-09-765-527-204	Sequence 204, App
21	33	100.0	10	9	US-09-765-527-205	Sequence 205, App
22	33	100.0	10	9	US-09-881-490-179	Sequence 179, App
23	33	100.0	10	9	US-09-881-490-180	Sequence 180, App
24	33	100.0	11	9	US-09-765-527-206	Sequence 206, App
25	33	100.0	11	9	US-09-881-490-181	Sequence 181, App
26	30	90.9	9	9	US-09-765-527-199	Sequence 199, App
27	30	90.9	9	9	US-09-881-490-167	Sequence 167, App
28	29	87.9	620	10	US-09-965-602-58	Sequence 58, Appl
29	29	87.9	1160	9	US-09-815-242-13913	Sequence 13913, A
30	28	84.8	5	9	US-09-765-527-186	Sequence 186, App
31	28	84.8	5	9	US-09-881-490-153	Sequence 153, App
32	28	84.8	5	9	US-09-881-490-209	Sequence 209, App
33	28	84.8	6	9	US-09-765-527-181	Sequence 181, App
34	28	84.8	6	9	US-09-765-527-182	Sequence 182, App
35	28	84.8	6	9	US-09-765-527-183	Sequence 183, App
36	28	84.8	6	9	US-09-881-490-148	Sequence 148, App
37	28	84.8	6	9	US-09-881-490-149	Sequence 149, App
38	28	84.8	6	9	US-09-881-490-211	Sequence 211, App
39	28	84.8	7	9	US-09-765-527-173	Sequence 173, App
40	28	84.8	7	9	US-09-765-527-174	Sequence 174, App
41	28	84.8	7	9	US-09-765-527-175	Sequence 175, App
42	28	84.8	7	9	US-09-765-527-191	Sequence 191, App
43	28	84.8	7	9	US-09-881-490-140	Sequence 140, App
44	28	84.8	7	9	US-09-881-490-141	Sequence 141, App
45	28	84.8	7	9	US-09-881-490-142	Sequence 142, App

ALIGNMENTS

RESULT 1
US-09-765-527-187
; Sequence 187, Application US/09765527
; Patent No. US20020006638A1
; GENERAL INFORMATION:
; APPLICANT: Better, Marc D.
; TITLE OF INVENTION: Methods for Recombinant Microbial Production of
; Fusion Proteins and BPI-Derived Peptides
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/765,527
; FILING DATE: 18-Jan-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/621,803
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Borun, Michael F.
; REGISTRATION NUMBER: 25,447
; REFERENCE/DOCKET NUMBER: 27129/33199
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 187:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids

```
;            TYPE: amino acid
;            TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        FEATURE:
;            NAME/KEY:  misc_feature
;            OTHER INFORMATION:  "XMP.321"
;        FEATURE:
;            NAME/KEY:  Modified-site
;            LOCATION:  C-Terminus
;            OTHER INFORMATION:  /label= Amidation
;            /note= "The C-Terminus is Amidated."
;        SEQUENCE DESCRIPTION: SEQ ID NO: 187:
US-09-765-527-187

  Query Match             100.0%;  Score 33;  DB 9;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 5.3e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              ||||||
Db          1 WLIQLK 6


RESULT 2
US-09-881-490-154
; Sequence 154, Application US/09881490
; Patent No. US20020077298A1
;   GENERAL INFORMATION:
;        APPLICANT: Little II, Roger G.
;                   Lim, Edward
;                   Fadem, Mitchell B.
;        TITLE OF INVENTION: Anti-Fungal Peptides
;        NUMBER OF SEQUENCES: 211
;        CORRESPONDENCE ADDRESS:
;            ADDRESSEE: McAndrews, Held & Malloy, Ltd.
;            STREET: 500 West Madison Street, 34th FloorDrive
;            CITY: Chicago
;            STATE: Illinois
;            COUNTRY: United States of America
;            ZIP: 60661
;        COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/881,490
;            FILING DATE: 14-Jun-2001
;        PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/119,858
;            FILING DATE: <Unknown>
;            APPLICATION NUMBER: 08/372,105
;            FILING DATE: 13-JAN-95
;            APPLICATION NUMBER: 08/306,473
;            FILING DATE: 15-SEP-94
;            APPLICATION NUMBER: 08/273,540
;            FILING DATE: 11-JUL-94
;            APPLICATION NUMBER: 08/209,762
;            FILING DATE: 11-MAR-94
;            APPLICATION NUMBER: 08/183,222
;            FILING DATE: 14-JAN-94
;            APPLICATION NUMBER: 08/093,202
;            FILING DATE: 15-JUL-93
;            APPLICATION NUMBER: 08/030,644
;            FILING DATE: 12-MAR-93
;        ATTORNEY/AGENT INFORMATION:
;            NAME: McNicholas, Janet M.
;            REGISTRATION NUMBER: 32,918
;            REFERENCE/DOCKET NUMBER: 100-238/11021US01
;        TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 312/707-8889
;            TELEFAX: 312/707-9155
;            TELEX: 650 388-1248
;   INFORMATION FOR SEQ ID NO: 154:
;        SEQUENCE CHARACTERISTICS:
;            LENGTH: 6 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        FEATURE:
;            NAME/KEY:  misc_feature
;            OTHER INFORMATION:  "XMP.321"
;        FEATURE:
;            NAME/KEY:  Modified-site
;            LOCATION:  C-Terminus
;            OTHER INFORMATION:  /label= Amidation
;            /note= "The C-Terminus is Amidated"
;        SEQUENCE DESCRIPTION: SEQ ID NO: 154:
US-09-881-490-154

  Query Match             100.0%;  Score 33;  DB 9;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 5.3e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              ||||||
Db          1 WLIQLK 6


RESULT 3
US-09-765-527-188
; Sequence 188, Application US/09765527
; Patent No. US20020006638A1
;   GENERAL INFORMATION:
;        APPLICANT: Better, Marc D.
;        TITLE OF INVENTION: Methods for Recombinant Microbial Production of
;                            Fusion Proteins and BPI-Derived Peptides
;        NUMBER OF SEQUENCES: 265
;        CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
;            STREET: 6300 Sears Tower, 233 South Wacker Drive
;            CITY: Chicago
;            STATE: Illinois
;            COUNTRY: United States of America
;            ZIP: 60606-6402
;        COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/765,527
;            FILING DATE: 18-Jan-2001
;        PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 08/621,803
;            FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;            NAME: Borun, Michael F.
;            REGISTRATION NUMBER: 25,447
;            REFERENCE/DOCKET NUMBER: 27129/33199
;        TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 312/474-6300
;            TELEFAX: 312/474-0448
;            TELEX: 25-3856
;   INFORMATION FOR SEQ ID NO: 188:
;        SEQUENCE CHARACTERISTICS:
;            LENGTH: 7 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        FEATURE:
;            NAME/KEY:  misc_feature
;            OTHER INFORMATION:  "XMP.322"
;        FEATURE:
;            NAME/KEY:  Modified-site
;            LOCATION:  C-Terminus
```

```
;             OTHER INFORMATION:  /label= Amidation
;             /note= "The C-Terminus is Amidated."
;         SEQUENCE DESCRIPTION: SEQ ID NO: 188:
US-09-765-527-188

  Query Match             100.0%;  Score 33;  DB 9;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 5.3e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              ||||||
Db          1 WLIQLK 6


RESULT 4
US-09-765-527-189
; Sequence 189, Application US/09765527
; Patent No. US20020006638A1
;   GENERAL INFORMATION:
;         APPLICANT: Better, Marc D.
;         TITLE OF INVENTION: Methods for Recombinant Microbial Production of
;                             Fusion Proteins and BPI-Derived Peptides
;         NUMBER OF SEQUENCES: 265
;         CORRESPONDENCE ADDRESS:
;              ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
;              STREET: 6300 Sears Tower, 233 South Wacker Drive
;              CITY: Chicago
;              STATE: Illinois
;              COUNTRY: United States of America
;              ZIP: 60606-6402
;         COMPUTER READABLE FORM:
;              MEDIUM TYPE: Floppy disk
;              COMPUTER: IBM PC compatible
;              OPERATING SYSTEM: PC-DOS/MS-DOS
;              SOFTWARE: PatentIn Release #1.0, Version #1.25
;         CURRENT APPLICATION DATA:
;              APPLICATION NUMBER: US/09/765,527
;              FILING DATE: 18-Jan-2001
;         PRIOR APPLICATION DATA:
;              APPLICATION NUMBER: 08/621,803
;              FILING DATE: <Unknown>
;         ATTORNEY/AGENT INFORMATION:
;              NAME: Borun, Michael F.
;              REGISTRATION NUMBER: 25,447
;              REFERENCE/DOCKET NUMBER: 27129/33199
;         TELECOMMUNICATION INFORMATION:
;              TELEPHONE: 312/474-6300
;              TELEFAX: 312/474-0448
;              TELEX: 25-3856
;    INFORMATION FOR SEQ ID NO: 189:
;         SEQUENCE CHARACTERISTICS:
;              LENGTH: 7 amino acids
;              TYPE: amino acid
;              TOPOLOGY: linear
;         MOLECULE TYPE: peptide
;         FEATURE:
;              NAME/KEY:  misc_feature
;              OTHER INFORMATION:  "XMP.323"
;         FEATURE:
;              NAME/KEY:  Modified-site
;              LOCATION:  C-Terminus
;              OTHER INFORMATION:  /label= Amidation
;              /note= "The C-Terminus is Amidated."
;         SEQUENCE DESCRIPTION: SEQ ID NO: 189:
US-09-765-527-189

  Query Match             100.0%;  Score 33;  DB 9;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 5.3e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              ||||||
Db          2 WLIQLK 7


RESULT 5
US-09-765-527-190
; Sequence 190, Application US/09765527
; Patent No. US20020006638A1
;   GENERAL INFORMATION:
;         APPLICANT: Better, Marc D.
;         TITLE OF INVENTION: Methods for Recombinant Microbial Production of
;                             Fusion Proteins and BPI-Derived Peptides
;         NUMBER OF SEQUENCES: 265
;         CORRESPONDENCE ADDRESS:
;              ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
;              STREET: 6300 Sears Tower, 233 South Wacker Drive
;              CITY: Chicago
;              STATE: Illinois
;              COUNTRY: United States of America
;              ZIP: 60606-6402
;         COMPUTER READABLE FORM:
;              MEDIUM TYPE: Floppy disk
;              COMPUTER: IBM PC compatible
;              OPERATING SYSTEM: PC-DOS/MS-DOS
;              SOFTWARE: PatentIn Release #1.0, Version #1.25
;         CURRENT APPLICATION DATA:
;              APPLICATION NUMBER: US/09/765,527
;              FILING DATE: 18-Jan-2001
;         PRIOR APPLICATION DATA:
;              APPLICATION NUMBER: 08/621,803
;              FILING DATE: <Unknown>
;         ATTORNEY/AGENT INFORMATION:
;              NAME: Borun, Michael F.
;              REGISTRATION NUMBER: 25,447
;              REFERENCE/DOCKET NUMBER: 27129/33199
;         TELECOMMUNICATION INFORMATION:
;              TELEPHONE: 312/474-6300
;              TELEFAX: 312/474-0448
;              TELEX: 25-3856
;    INFORMATION FOR SEQ ID NO: 190:
;         SEQUENCE CHARACTERISTICS:
;              LENGTH: 7 amino acids
;              TYPE: amino acid
;              TOPOLOGY: linear
;         MOLECULE TYPE: peptide
;         FEATURE:
;              NAME/KEY:  misc_feature
;              OTHER INFORMATION:  "XMP.324"
;         FEATURE:
;              NAME/KEY:  Modified-site
;              LOCATION:  C-Terminus
;              OTHER INFORMATION:  /label= Amidation
;              /note= "The C-Terminus is Amidated."
;         SEQUENCE DESCRIPTION: SEQ ID NO: 190:
US-09-765-527-190

  Query Match             100.0%;  Score 33;  DB 9;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 5.3e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              ||||||
Db          2 WLIQLK 7


RESULT 6
US-09-881-490-155
; Sequence 155, Application US/09881490
; Patent No. US20020077298A1
;   GENERAL INFORMATION:
;         APPLICANT: Little II, Roger G.
;                    Lim, Edward
;                    Fadem, Mitchell B.
;         TITLE OF INVENTION: Anti-Fungal Peptides
;         NUMBER OF SEQUENCES: 211
```

```
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: McAndrews, Held & Malloy, Ltd.
;            STREET: 500 West Madison Street, 34th FloorDrive
;            CITY: Chicago
;            STATE: Illinois
;            COUNTRY: United States of America
;            ZIP: 60661
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/881,490
;            FILING DATE: 14-Jun-2001
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/119,858
;            FILING DATE: <Unknown>
;            APPLICATION NUMBER: 08/372,105
;            FILING DATE: 13-JAN-95
;            APPLICATION NUMBER: 08/306,473
;            FILING DATE: 15-SEP-94
;            APPLICATION NUMBER: 08/273,540
;            FILING DATE: 11-JUL-94
;            APPLICATION NUMBER: 08/209,762
;            FILING DATE: 11-MAR-94
;            APPLICATION NUMBER: 08/183,222
;            FILING DATE: 14-JAN-94
;            APPLICATION NUMBER: 08/093,202
;            FILING DATE: 15-JUL-93
;            APPLICATION NUMBER: 08/030,644
;            FILING DATE: 12-MAR-93
;       ATTORNEY/AGENT INFORMATION:
;            NAME: McNicholas, Janet M.
;            REGISTRATION NUMBER: 32,918
;            REFERENCE/DOCKET NUMBER: 100-238/11021US01
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 312/707-8889
;            TELEFAX: 312/707-9155
;            TELEX: 650 388-1248
;   INFORMATION FOR SEQ ID NO: 155:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 7 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: peptide
;       FEATURE:
;            NAME/KEY:  misc_feature
;            OTHER INFORMATION:  "XMP.322"
;       FEATURE:
;            NAME/KEY:  Modified-site
;            LOCATION:  C-Terminus
;            OTHER INFORMATION:  /label= Amidation
;            /note= "The C-Terminus is Amidated"
;       SEQUENCE DESCRIPTION: SEQ ID NO: 155:
US-09-881-490-155

  Query Match             100.0%;  Score 33;  DB 9;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              ||||||
Db          1 WLIQLK 6


RESULT 7
US-09-881-490-156
; Sequence 156, Application US/09881490
; Patent No. US20020077298A1
;   GENERAL INFORMATION:
;       APPLICANT: Little II, Roger G.
;                  Lim, Edward
;                  Fadem, Mitchell B.
;       TITLE OF INVENTION: Anti-Fungal Peptides
;       NUMBER OF SEQUENCES: 211
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: McAndrews, Held & Malloy, Ltd.
;            STREET: 500 West Madison Street, 34th FloorDrive
;            CITY: Chicago
;            STATE: Illinois
;            COUNTRY: United States of America
;            ZIP: 60661
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/881,490
;            FILING DATE: 14-Jun-2001
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/119,858
;            FILING DATE: <Unknown>
;            APPLICATION NUMBER: 08/372,105
;            FILING DATE: 13-JAN-95
;            APPLICATION NUMBER: 08/306,473
;            FILING DATE: 15-SEP-94
;            APPLICATION NUMBER: 08/273,540
;            FILING DATE: 11-JUL-94
;            APPLICATION NUMBER: 08/209,762
;            FILING DATE: 11-MAR-94
;            APPLICATION NUMBER: 08/183,222
;            FILING DATE: 14-JAN-94
;            APPLICATION NUMBER: 08/093,202
;            FILING DATE: 15-JUL-93
;            APPLICATION NUMBER: 08/030,644
;            FILING DATE: 12-MAR-93
;       ATTORNEY/AGENT INFORMATION:
;            NAME: McNicholas, Janet M.
;            REGISTRATION NUMBER: 32,918
;            REFERENCE/DOCKET NUMBER: 100-238/11021US01
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 312/707-8889
;            TELEFAX: 312/707-9155
;            TELEX: 650 388-1248
;   INFORMATION FOR SEQ ID NO: 156:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 7 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: peptide
;       FEATURE:
;            NAME/KEY:  misc_feature
;            OTHER INFORMATION:  "XMP.323"
;       FEATURE:
;            NAME/KEY:  Modified-site
;            LOCATION:  C-Terminus
;            OTHER INFORMATION:  /label= Amidation
;            /note= "The C-Terminus is Amidated"
;       SEQUENCE DESCRIPTION: SEQ ID NO: 156:
US-09-881-490-156

  Query Match             100.0%;  Score 33;  DB 9;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 5.3e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              ||||||
Db          2 WLIQLK 7


RESULT 8
US-09-765-527-192
; Sequence 192, Application US/09765527
; Patent No. US20020006638A1
```

```
;   GENERAL INFORMATION:
;        APPLICANT: Better, Marc D.
;        TITLE OF INVENTION: Methods for Recombinant Microbial Production of
;                            Fusion Proteins and BPI-Derived Peptides
;        NUMBER OF SEQUENCES: 265
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
;             STREET: 6300 Sears Tower, 233 South Wacker Drive
;             CITY: Chicago
;             STATE: Illinois
;             COUNTRY: United States of America
;             ZIP: 60606-6402
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/765,527
;             FILING DATE: 18-Jan-2001
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/621,803
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Borun, Michael F.
;             REGISTRATION NUMBER: 25,447
;             REFERENCE/DOCKET NUMBER: 27129/33199
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 312/474-6300
;             TELEFAX: 312/474-0448
;             TELEX: 25-3856
;   INFORMATION FOR SEQ ID NO: 192:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 8 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        FEATURE:
;             NAME/KEY:  misc_feature
;             OTHER INFORMATION:  "XMP.326"
;        FEATURE:
;             NAME/KEY:  Modified-site
;             LOCATION:  C-Terminus
;             OTHER INFORMATION:  /label= Amidation
;             /note= "The C-Terminus is Amidated."
;        SEQUENCE DESCRIPTION: SEQ ID NO: 192:
US-09-765-527-192

  Query Match             100.0%;  Score 33;  DB 9;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 5.3e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              ||||||
Db          3 WLIQLK 8


RESULT 9
US-09-881-490-157
; Sequence 157, Application US/09881490
; Patent No. US20020077298A1
;   GENERAL INFORMATION:
;        APPLICANT: Little II, Roger G.
;                   Lim, Edward
;                   Fadem, Mitchell B.
;        TITLE OF INVENTION: Anti-Fungal Peptides
;        NUMBER OF SEQUENCES: 211
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: McAndrews, Held & Malloy, Ltd.
;             STREET: 500 West Madison Street, 34th FloorDrive
;             CITY: Chicago
;             STATE: Illinois
;             COUNTRY: United States of America
;             ZIP: 60661
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/881,490
;             FILING DATE: 14-Jun-2001
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/119,858
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: 08/372,105
;             FILING DATE: 13-JAN-95
;             APPLICATION NUMBER: 08/306,473
;             FILING DATE: 15-SEP-94
;             APPLICATION NUMBER: 08/273,540
;             FILING DATE: 11-JUL-94
;             APPLICATION NUMBER: 08/209,762
;             FILING DATE: 11-MAR-94
;             APPLICATION NUMBER: 08/183,222
;             FILING DATE: 14-JAN-94
;             APPLICATION NUMBER: 08/093,202
;             FILING DATE: 15-JUL-93
;             APPLICATION NUMBER: 08/030,644
;             FILING DATE: 12-MAR-93
;        ATTORNEY/AGENT INFORMATION:
;             NAME: McNicholas, Janet M.
;             REGISTRATION NUMBER: 32,918
;             REFERENCE/DOCKET NUMBER: 100-238/110210US01
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 312/707-8889
;             TELEFAX: 312/707-9155
;             TELEX: 650 388-1248
;   INFORMATION FOR SEQ ID NO: 157:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 8 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        FEATURE:
;             NAME/KEY:  misc_feature
;             OTHER INFORMATION:  "XMP.324"
;        FEATURE:
;             NAME/KEY:  Modified-site
;             LOCATION:  C-Terminus
;             OTHER INFORMATION:  /label= Amidation
;             /note= "The C-Terminus is Amidated"
;        SEQUENCE DESCRIPTION: SEQ ID NO: 157:
US-09-881-490-157

  Query Match             100.0%;  Score 33;  DB 9;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 5.3e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              ||||||
Db          2 WLIQLK 7


RESULT 10
US-09-881-490-159
; Sequence 159, Application US/09881490
; Patent No. US20020077298A1
;   GENERAL INFORMATION:
;        APPLICANT: Little II, Roger G.
;                   Lim, Edward
;                   Fadem, Mitchell B.
;        TITLE OF INVENTION: Anti-Fungal Peptides
;        NUMBER OF SEQUENCES: 211
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: McAndrews, Held & Malloy, Ltd.
;             STREET: 500 West Madison Street, 34th FloorDrive
```

```
;             CITY: Chicago
;             STATE: Illinois
;             COUNTRY: United States of America
;             ZIP: 60661
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/881,490
;             FILING DATE: 14-Jun-2001
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/119,858
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: 08/372,105
;             FILING DATE: 13-JAN-95
;             APPLICATION NUMBER: 08/306,473
;             FILING DATE: 15-SEP-94
;             APPLICATION NUMBER: 08/273,540
;             FILING DATE: 11-JUL-94
;             APPLICATION NUMBER: 08/209,762
;             FILING DATE: 11-MAR-94
;             APPLICATION NUMBER: 08/183,222
;             FILING DATE: 14-JAN-94
;             APPLICATION NUMBER: 08/093,202
;             FILING DATE: 15-JUL-93
;             APPLICATION NUMBER: 08/030,644
;             FILING DATE: 12-MAR-93
;        ATTORNEY/AGENT INFORMATION:
;             NAME: McNicholas, Janet M.
;             REGISTRATION NUMBER: 32,918
;             REFERENCE/DOCKET NUMBER: 100-238/11021US01
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 312/707-8889
;             TELEFAX: 312/707-9155
;             TELEX: 650 388-1248
;   INFORMATION FOR SEQ ID NO: 159:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 8 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        FEATURE:
;             NAME/KEY:  misc_feature
;             OTHER INFORMATION:  "XMP.326"
;        FEATURE:
;             NAME/KEY:  Modified-site
;             LOCATION:  C-Terminus
;             OTHER INFORMATION:  /label= Amidation
;             /note= "The C-Terminus is Amidated"
;        SEQUENCE DESCRIPTION: SEQ ID NO: 159:
US-09-881-490-159

  Query Match             100.0%;  Score 33;  DB 9;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 5.3e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              ||||||
Db          3 WLIQLK 8


RESULT 11
US-09-765-527-193
; Sequence 193, Application US/09765527
; Patent No. US20020006638A1
;   GENERAL INFORMATION:
;        APPLICANT: Better, Marc D.
;        TITLE OF INVENTION: Methods for Recombinant Microbial Production of
;                            Fusion Proteins and BPI-Derived Peptides
;        NUMBER OF SEQUENCES: 265
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
;             STREET: 6300 Sears Tower, 233 South Wacker Drive
;             CITY: Chicago
;             STATE: Illinois
;             COUNTRY: United States of America
;             ZIP: 60606-6402
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/765,527
;             FILING DATE: 18-Jan-2001
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/621,803
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Borun, Michael F.
;             REGISTRATION NUMBER: 25,447
;             REFERENCE/DOCKET NUMBER: 27129/33199
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 312/474-6300
;             TELEFAX: 312/474-0448
;             TELEX: 25-3856
;   INFORMATION FOR SEQ ID NO: 193:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 9 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        FEATURE:
;             NAME/KEY:  misc_feature
;             OTHER INFORMATION:  "XMP.327"
;        FEATURE:
;             NAME/KEY:  Modified-site
;             LOCATION:  C-Terminus
;             OTHER INFORMATION:  /label= Amidation
;             /note= "The C-Terminus is Amidated."
;        SEQUENCE DESCRIPTION: SEQ ID NO: 193:
US-09-765-527-193

  Query Match             100.0%;  Score 33;  DB 9;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 5.3e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              ||||||
Db          3 WLIQLK 8


RESULT 12
US-09-765-527-197
; Sequence 197, Application US/09765527
; Patent No. US20020006638A1
;   GENERAL INFORMATION:
;        APPLICANT: Better, Marc D.
;        TITLE OF INVENTION: Methods for Recombinant Microbial Production of
;                            Fusion Proteins and BPI-Derived Peptides
;        NUMBER OF SEQUENCES: 265
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
;             STREET: 6300 Sears Tower, 233 South Wacker Drive
;             CITY: Chicago
;             STATE: Illinois
;             COUNTRY: United States of America
;             ZIP: 60606-6402
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
```

```
;             APPLICATION NUMBER: US/09/765,527
;             FILING DATE: 18-Jan-2001
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/621,803
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Borun, Michael F.
;             REGISTRATION NUMBER: 25,447
;             REFERENCE/DOCKET NUMBER: 27129/33199
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 312/474-6300
;             TELEFAX: 312/474-0448
;             TELEX: 25-3856
;   INFORMATION FOR SEQ ID NO: 197:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 9 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        FEATURE:
;             NAME/KEY:  misc_feature
;             OTHER INFORMATION:  "XMP.331"
;        FEATURE:
;             NAME/KEY:  Modified-site
;             LOCATION:  1
;             OTHER INFORMATION:  /label= Acetylated
;             /note= "Position 1 is acetylated."
;        FEATURE:
;             NAME/KEY:  Modified-site
;             LOCATION:  C-Terminus
;             OTHER INFORMATION:  /label= Amidation
;             /note= "The C-Terminus is Amidated."
;        SEQUENCE DESCRIPTION: SEQ ID NO: 197:
US-09-765-527-197

  Query Match             100.0%;  Score 33;  DB 9;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 5.3e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              ||||||
Db          3 WLIQLK 8


RESULT 13
US-09-765-527-198
; Sequence 198, Application US/09765527
; Patent No. US20020006638A1
;   GENERAL INFORMATION:
;        APPLICANT: Better, Marc D.
;        TITLE OF INVENTION: Methods for Recombinant Microbial Production of
;                            Fusion Proteins and BPI-Derived Peptides
;        NUMBER OF SEQUENCES: 265
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
;             STREET: 6300 Sears Tower, 233 South Wacker Drive
;             CITY: Chicago
;             STATE: Illinois
;             COUNTRY: United States of America
;             ZIP: 60606-6402
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/765,527
;             FILING DATE: 18-Jan-2001
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/621,803
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Borun, Michael F.
;             REGISTRATION NUMBER: 25,447
;             REFERENCE/DOCKET NUMBER: 27129/33199
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 312/474-6300
;             TELEFAX: 312/474-0448
;             TELEX: 25-3856
;   INFORMATION FOR SEQ ID NO: 198:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 9 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        FEATURE:
;             NAME/KEY:  misc_feature
;             OTHER INFORMATION:  "XMP.335"
;        FEATURE:
;             NAME/KEY:  Modified-site
;             LOCATION:  C-Terminus
;             OTHER INFORMATION:  /label= Amidation
;             /note= "The C-Terminus is Amidated."
;        SEQUENCE DESCRIPTION: SEQ ID NO: 198:
US-09-765-527-198

  Query Match             100.0%;  Score 33;  DB 9;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 5.3e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              ||||||
Db          3 WLIQLK 8


RESULT 14
US-09-881-490-160
; Sequence 160, Application US/09881490
; Patent No. US20020077298A1
;   GENERAL INFORMATION:
;        APPLICANT: Little II, Roger G.
;                   Lim, Edward
;                   Fadem, Mitchell B.
;        TITLE OF INVENTION: Anti-Fungal Peptides
;        NUMBER OF SEQUENCES: 211
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: McAndrews, Held & Malloy, Ltd.
;             STREET: 500 West Madison Street, 34th FloorDrive
;             CITY: Chicago
;             STATE: Illinois
;             COUNTRY: United States of America
;             ZIP: 60661
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/881,490
;             FILING DATE: 14-Jun-2001
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/119,858
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: 08/372,105
;             FILING DATE: 13-JAN-95
;             APPLICATION NUMBER: 08/306,473
;             FILING DATE: 15-SEP-94
;             APPLICATION NUMBER: 08/273,540
;             FILING DATE: 11-JUL-94
;             APPLICATION NUMBER: 08/209,762
;             FILING DATE: 11-MAR-94
;             APPLICATION NUMBER: 08/183,222
;             FILING DATE: 14-JAN-94
;             APPLICATION NUMBER: 08/093,202
;             FILING DATE: 15-JUL-93
;             APPLICATION NUMBER: 08/030,644
```

```
;               FILING DATE: 12-MAR-93
;          ATTORNEY/AGENT INFORMATION:
;               NAME: McNicholas, Janet M.
;               REGISTRATION NUMBER: 32,918
;               REFERENCE/DOCKET NUMBER: 100-238/11021US01
;          TELECOMMUNICATION INFORMATION:
;               TELEPHONE: 312/707-8889
;               TELEFAX: 312/707-9155
;               TELEX: 650 388-1248
;     INFORMATION FOR SEQ ID NO: 160:
;          SEQUENCE CHARACTERISTICS:
;               LENGTH: 9 amino acids
;               TYPE: amino acid
;               TOPOLOGY: linear
;          MOLECULE TYPE: peptide
;          FEATURE:
;               NAME/KEY:  misc_feature
;               OTHER INFORMATION:  "XMP.327"
;          FEATURE:
;               NAME/KEY:  Modified-site
;               LOCATION:  C-Terminus
;               OTHER INFORMATION:  /label= Amidation
;               /note= "The C-Terminus is Amidated"
;          SEQUENCE DESCRIPTION: SEQ ID NO: 160:
US-09-881-490-160

  Query Match             100.0%;  Score 33;  DB 9;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 5.3e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              ||||||
Db          3 WLIQLK 8


RESULT 15
US-09-881-490-162
; Sequence 162, Application US/09881490
; Patent No. US20020077298A1
;   GENERAL INFORMATION:
;          APPLICANT: Little II, Roger G.
;                     Lim, Edward
;                     Fadem, Mitchell B.
;          TITLE OF INVENTION: Anti-Fungal Peptides
;          NUMBER OF SEQUENCES: 211
;          CORRESPONDENCE ADDRESS:
;               ADDRESSEE: McAndrews, Held & Malloy, Ltd.
;               STREET: 500 West Madison Street, 34th FloorDrive
;               CITY: Chicago
;               STATE: Illinois
;               COUNTRY: United States of America
;               ZIP: 60661
;          COMPUTER READABLE FORM:
;               MEDIUM TYPE: Floppy disk
;               COMPUTER: IBM PC compatible
;               OPERATING SYSTEM: PC-DOS/MS-DOS
;               SOFTWARE: PatentIn Release #1.0, Version #1.25
;          CURRENT APPLICATION DATA:
;               APPLICATION NUMBER: US/09/881,490
;               FILING DATE: 14-Jun-2001
;          PRIOR APPLICATION DATA:
;               APPLICATION NUMBER: 09/119,858
;               FILING DATE: <Unknown>
;               APPLICATION NUMBER: 08/372,105
;               FILING DATE: 13-JAN-95
;               APPLICATION NUMBER: 08/306,473
;               FILING DATE: 15-SEP-94
;               APPLICATION NUMBER: 08/273,540
;               FILING DATE: 11-JUL-94
;               APPLICATION NUMBER: 08/209,762
;               FILING DATE: 11-MAR-94
;               APPLICATION NUMBER: 08/183,222
;               FILING DATE: 14-JAN-94
;               APPLICATION NUMBER: 08/093,202
;               FILING DATE: 15-JUL-93
;               APPLICATION NUMBER: 08/030,644
;               FILING DATE: 12-MAR-93
;          ATTORNEY/AGENT INFORMATION:
;               NAME: McNicholas, Janet M.
;               REGISTRATION NUMBER: 32,918
;               REFERENCE/DOCKET NUMBER: 100-238/11021US01
;          TELECOMMUNICATION INFORMATION:
;               TELEPHONE: 312/707-8889
;               TELEFAX: 312/707-9155
;               TELEX: 650 388-1248
;     INFORMATION FOR SEQ ID NO: 162:
;          SEQUENCE CHARACTERISTICS:
;               LENGTH: 9 amino acids
;               TYPE: amino acid
;               TOPOLOGY: linear
;          MOLECULE TYPE: peptide
;          FEATURE:
;               NAME/KEY:  misc_feature
;               OTHER INFORMATION:  "XMP.331"
;          FEATURE:
;               NAME/KEY:  Modified-site
;               LOCATION:  1
;               OTHER INFORMATION:  /label= Acetylated
;               /note= "Position 1 is acetylated"
;          FEATURE:
;               NAME/KEY:  Modified-site
;               LOCATION:  C-Terminus
;               OTHER INFORMATION:  /label= Amidation
;               /note= "The C-Terminus is Amidated"
;          SEQUENCE DESCRIPTION: SEQ ID NO: 162:
US-09-881-490-162

  Query Match             100.0%;  Score 33;  DB 9;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 5.3e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              ||||||
Db          3 WLIQLK 8


Search completed: October 14, 2003, 14:44:49
Job time : 14.625 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 14:33:18 ; Search time 7.875 Seconds
(without alignments)
73.271 Million cell updates/sec

Title: US-09-881-490-154
Perfect score: 33
Sequence: 1 WLIQLK 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	32	97.0	182	2	A90193	conserved hypothet
2	29	87.9	207	2	G84983	5-amino-6-(5-phosp
3	29	87.9	287	2	B90540	conserved hypothet
4	29	87.9	341	1	E64444	hypothetical prote
5	29	87.9	388	1	WMLJBT	bel-2 protein - si
6	29	87.9	390	2	T38370	hypothetical prote
7	29	87.9	416	2	T02194	probable pectinace
8	29	87.9	420	2	T47998	pectinacetylestera
9	29	87.9	422	2	T01197	pectin acetylester
10	29	87.9	620	2	S56790	probable membrane
11	29	87.9	756	2	G86308	Similar to disease
12	29	87.9	1160	1	A45915	DNA-directed DNA p
13	29	87.9	1160	2	AI0530	DNA polymerase III
14	28	84.8	59	2	AE2023	hypothetical prote
15	28	84.8	62	2	A86805	hypothetical prote
16	28	84.8	146	2	B86721	transcription regu
17	28	84.8	162	2	T02072	probable glycine c
18	28	84.8	166	2	I37052	H+-transporting tw
19	28	84.8	179	2	T49508	hypothetical prote
20	28	84.8	180	1	GCHU	glucagon precursor
21	28	84.8	180	1	GCGP	glucagon precursor
22	28	84.8	180	1	GCRTDU	glucagon precursor
23	28	84.8	180	1	GCRT	glucagon precursor
24	28	84.8	180	1	GCHY	glucagon precursor
25	28	84.8	180	2	A57294	glucagon precursor
26	28	84.8	189	2	AH1029	probable membrane
27	28	84.8	194	2	PC1136	amidophosphoribosy
28	28	84.8	207	2	T32983	hypothetical prote
29	28	84.8	215	2	C70471	hypothetical prote
30	28	84.8	215	2	AE2177	hypothetical prote
31	28	84.8	226	2	T11838	H+-transporting tw
32	28	84.8	226	2	T11498	H+-transporting tw
33	28	84.8	226	2	S26156	H+-transporting tw
34	28	84.8	226	2	S41840	H+-transporting tw
35	28	84.8	226	2	T10977	H+-transporting tw
36	28	84.8	226	2	F59153	H+-transporting tw
37	28	84.8	230	2	T45555	H+-transporting tw
38	28	84.8	250	4	S43383	bactericidal/perme
39	28	84.8	275	2	H81074	phytoene synthase-
40	28	84.8	276	2	D81868	probable transfera
41	28	84.8	292	2	A64624	hypothetical prote
42	28	84.8	326	2	T26368	hypothetical prote
43	28	84.8	368	2	T47894	actin-like protein
44	28	84.8	379	2	AG3355	membrane metallopr
45	28	84.8	394	2	S76353	probable aspartate

ALIGNMENTS

RESULT 1
A90193
conserved hypothetical protein [imported] - Sulfolobus solfataricus
C;Species: Sulfolobus solfataricus
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
C;Accession: A90193
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redd arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A;Description: Sulfolobus solfataricus complete genome.
A;Reference number: A99139
A;Accession: A90193
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-182 <KUR>
A;Cross-references: GB:AE006641; NID:g13813634; PIDN:AAK40800.1; GSPDB:GN00155
C;Genetics:
A;Gene: SSO0479

Query Match 97.0%; Score 32; DB 2; Length 182;
Best Local Similarity 83.3%; Pred. No. 13;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 WLIQLK 6
              ||:|||
Db        110 WLVQLK 115
```

RESULT 2
G84983
5-amino-6-(5-phosphoribosylamino)uracil reductase (EC 1.1.1.193) [imported] - Buchn
N;Alternate names: riboflavin reductase
C;Species: Buchnera sp.
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001
C;Accession: G84983
R;Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.
Nature 407, 81-86, 2000
A;Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera
A;Reference number: A84930; MUID:20445173; PMID:10993077
A;Accession: G84983
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-207 <STO>
A;Cross-references: GB:AP000398; GSPDB:GN00144
A;Experimental source: strain APS
C;Genetics:
A;Gene: ribD2; BU462
C;Keywords: oxidoreductase

Query Match 87.9%; Score 29; DB 2; Length 207;
Best Local Similarity 83.3%; Pred. No. 66;

```
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              |||:||
Db         90 WLIRLK 95


RESULT 3
B90540
conserved hypothetical protein MYPU_2260 [imported] - Mycoplasma pulmonis (strain UAB CT
C;Species: Mycoplasma pulmonis
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 03-Aug-2001
C;Accession: B90540
R;Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I.;
Nucleic Acids Res. 29, 2145-2153, 2001
A;Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulm
A;Reference number: A99512; MUID:21267165; PMID:11353084
A;Accession: B90540
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-287 <KUR>
A;Cross-references: GB:AL445566; PID:g14089639; PIDN:CAC13399.1; GSPDB:GN00153
A;Experimental source: strain UAB CTIP
C;Genetics:
A;Gene: MYPU_2260
A;Genetic code: SGC3

  Query Match             87.9%;  Score 29;  DB 2;  Length 287;
  Best Local Similarity   83.3%;  Pred. No. 92;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              |||:||
Db        272 WLIKLK 277


RESULT 4
E64444
hypothetical protein MJ1157 - Methanococcus jannaschii
C;Species: Methanococcus jannaschii
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
C;Accession: E64444
R;Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,
; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A;Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C
A;Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
A;Reference number: A64300; MUID:96337999; PMID:8688087
A;Accession: E64444
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-341 <BUL>
A;Cross-references: GB:U67558; GB:L77117; NID:g1591786; PIDN:AAB99161.1; PID:g1592323; I
C;Genetics:
A;Map position: FOR1097616-1098641
C;Superfamily: conserved hypothetical protein MJ1157

  Query Match             87.9%;  Score 29;  DB 1;  Length 341;
  Best Local Similarity   83.3%;  Pred. No. 1.1e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              |||:||
Db        333 WLIKLK 338


RESULT 5
WMLJBT
bel-2 protein - simian foamy virus (type 3, strain LK3)
C;Species: simian foamy virus
A;Note: host (African green monkey)
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 08-Apr-1994
C;Accession: E40820
R;Renne, R.; Friedl, E.; Schweizer, M.; Fleps, U.; Turek, R.; Neumann-Haefelin, D.
Virology 186, 597-608, 1992
A;Title: Genomic organization and expression of simian foamy virus type 3 (SFV-3).
A;Reference number: A40820; MUID:92124734; PMID:1310187
A;Accession: E40820
A;Molecule type: DNA
A;Residues: 1-388 <REN>
A;Cross-references: GB:M74895
C;Genetics:
A;Gene: bel-2
C;Superfamily: foamy virus bel-2 protein

  Query Match             87.9%;  Score 29;  DB 1;  Length 388;
  Best Local Similarity   83.3%;  Pred. No. 1.2e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              |||:||
Db        182 WLIRLK 187


RESULT 6
T38370
hypothetical protein SPAC25A8.02 - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 04-Mar-2000
C;Accession: T38370
R;Skelton, J.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, February 1998
A;Reference number: Z21789
A;Accession: T38370
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-390 <SKE>
A;Cross-references: EMBL:AL021809; PIDN:CAA16952.1; GSPDB:GN00066; SPDB:SPAC25A8.02
A;Experimental source: strain 972h-; cosmid c25A8
C;Genetics:
A;Gene: SPDB:SPAC25A8.02
A;Map position: 1
C;Superfamily: Schizosaccharomyces pombe hypothetical protein SPAC25A8.02

  Query Match             87.9%;  Score 29;  DB 2;  Length 390;
  Best Local Similarity   66.7%;  Pred. No. 1.2e+02;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              ||:||:
Db        333 WLVQLR 338


RESULT 7
T02194
probable pectinacetylesterase At2g46930 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 02-Feb-2001
C;Accession: T02194; A84909
R;Rounsley, S.D.; Lin, X.; Kaul, S.; Shea, T.P.; Fujii, C.Y.; Mason, T.M.; Shen, M.;
submitted to the EMBL Data Library, September 1998
A;Description: Arabidopsis thaliana chromosome II BAC F14M4 genomic sequence.
A;Reference number: Z14609
A;Accession: T02194
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-416 <ROU>
A;Cross-references: EMBL:AC004411; NID:g3522932; PID:g3522956
A;Experimental source: cultivar Columbia
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallor
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Vente
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
```

```
A;Accession: AB4909
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-416 <STO>
A;Cross-references: GB:AE002093; NID:g3522956; PIDN:AAC34238.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g46930; F14M4.24
A;Map position: 2
A;Introns: 67/1; 94/3; 138/1; 162/3; 191/3; 233/2; 253/3; 271/3; 297/3; 333/1; 381/3

  Query Match            87.9%;  Score 29;  DB 2;  Length 416;
  Best Local Similarity  83.3%;  Pred. No. 1.3e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              |||||:
Db         89 WLIQLE 94


RESULT 8
T47998
pectinacetylesterase-like protein T17J13.20 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 02-Sep-2000
C;Accession: T47998
R;Rieger, M.; Mueller-Auer, S.; Zipp, M.; Schaefer, M.; Mewes, H.W.; Lemcke, K.; Mayer,
submitted to the Protein Sequence Database, February 2000
A;Reference number: Z24482
A;Accession: T47998
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-420 <RIE>
A;Cross-references: EMBL:AL138651
A;Experimental source: cultivar Columbia; BAC clone T17J13
C;Genetics:
A;Map position: 3
A;Introns: 69/1; 96/3; 140/1; 164/3; 202/1; 236/2; 256/3; 274/3; 300/3; 336/1; 384/3
A;Note: T17J13.20

  Query Match            87.9%;  Score 29;  DB 2;  Length 420;
  Best Local Similarity  83.3%;  Pred. No. 1.3e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              |||||:
Db         91 WLIQLE 96


RESULT 9
T01197
pectin acetylesterase homolog F21E10.11 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 22-Oct-1999
C;Accession: T01197
R;Davidson, S.; Rohlfing, T.; David, M.; O'Brian, D.
submitted to the EMBL Data Library, April 1998
A;Description: The sequence of A. thaliana F21E10.
A;Reference number: Z14258
A;Accession: T01197
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-422 <DAV>
A;Cross-references: EMBL:AF058914; NID:g3047074; PID:g3047082; GSPDB:GN00063; ATSP:F21E1
A;Experimental source: cultivar Columbia
C;Genetics:
A;Gene: ATSP:F21E10.11
A;Map position: 5
A;Introns: 67/1; 94/3; 138/1; 162/3; 191/3; 233/2; 252/3; 277/3; 303/3; 339/1; 387/3

  Query Match            87.9%;  Score 29;  DB 2;  Length 422;
  Best Local Similarity  83.3%;  Pred. No. 1.3e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              |||||:
Db         89 WLIQLE 94


RESULT 10
S56790
probable membrane protein YJL019w - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein J1310
C;Species: Saccharomyces cerevisiae
C;Date: 08-Jul-1995 #sequence_revision 08-Sep-1995 #text_change 19-Apr-2002
C;Accession: S56790
R;To Van, D.; Perea, J.; Jacq, C.
submitted to the Protein Sequence Database, September 1995
A;Reference number: S56776
A;Accession: S56790
A;Molecule type: DNA
A;Residues: 1-620 <DEH>
A;Cross-references: EMBL:Z49294; NID:g1006737; PID:g1006738; GSPDB:GN00010; MIPS:YJL
C;Genetics:
A;Gene: MIPS:YJL019w
A;Cross-references: SGD:S0003556
A;Map position: 10L
C;Superfamily: Saccharomyces cerevisiae probable membrane protein YJL019w
C;Keywords: transmembrane protein

  Query Match            87.9%;  Score 29;  DB 2;  Length 620;
  Best Local Similarity  83.3%;  Pred. No. 2e+02;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              | ||||
Db        579 WFIQLK 584


RESULT 11
G86308
Similar to disease resistance proteins [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: G86308
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alc
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzi
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tall
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: G86308
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-756 <STO>
A;Cross-references: GB:AE005172; NID:g5734745; PIDN:AAD50010.1; GSPDB:GN00141
C;Genetics:
A;Map position: 1

  Query Match            87.9%;  Score 29;  DB 2;  Length 756;
  Best Local Similarity  83.3%;  Pred. No. 2.4e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              |||:||
Db        494 WLIKLK 499


RESULT 12
A45915
DNA-directed DNA polymerase (EC 2.7.7.7) III alpha chain - Salmonella typhimurium
C;Species: Salmonella typhimurium
```

C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 19-Jul-2002
C;Accession: A45915
R;Lancy, E.D.; Lifsics, M.R.; Munson, P.; Maurer, R.
J. Bacteriol. 171, 5581-5586, 1989
A;Title: Nucleotide sequence of dnaE, the gene for the polymerase subunit of DNA polymer
ubunit.
A;Reference number: A45915; MUID:90008797; PMID:2676978
A;Accession: A45915
A;Status: preliminary; translation not shown
A;Molecule type: DNA
A;Residues: 1-1160 <LAN>
A;Cross-references: GB:M29701; NID:g153951; PIDN:AAA27057.1; PID:g153952
C;Superfamily: DNA-directed DNA polymerase III alpha chain
C;Keywords: hydrolase; nucleotidyltransferase

 Query Match 87.9%; Score 29; DB 1; Length 1160;
 Best Local Similarity 66.7%; Pred. No. 3.7e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WLIQLK 6
||::||
Db 120 WLVELK 125

RESULT 13
AI0530
DNA polymerase III, alpha chain [imported] - Salmonella enterica subsp. enterica serovar
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AI0530
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AI0530
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1160 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD08689.1; PID:g16501512; GSPDB:GN00176
C;Genetics:
A;Gene: STY0254
C;Superfamily: DNA-directed DNA polymerase III alpha chain

 Query Match 87.9%; Score 29; DB 2; Length 1160;
 Best Local Similarity 66.7%; Pred. No. 3.7e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WLIQLK 6
||::||
Db 120 WLVELK 125

RESULT 14
AE2023
hypothetical protein asl1739 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C;Accession: AE2023
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AE2023
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-59 <KUR>
A;Cross-references: GB:BA000019; PIDN:BAB73438.1; PID:g17130829; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: asl1739

 Query Match 84.8%; Score 28; DB 2; Length 59;
 Best Local Similarity 66.7%; Pred. No. 31;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WLIQLK 6
||::||
Db 49 WLVKLK 54

RESULT 15
A86805
hypothetical protein yogE [imported] - Lactococcus lactis subsp. lactis (strain IL14
C;Species: Lactococcus lactis subsp. lactis
C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C;Accession: A86805
R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; E
Genome Res. 11, 731-753, 2001
A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lacti
A;Reference number: A86625; MUID:21235186; PMID:11337471
A;Accession: A86805
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-62 <STO>
A;Cross-references: GB:AE005176; PID:g12724431; PIDN:AAK05539.1; GSPDB:GN00146
A;Experimental source: strain IL1403
C;Genetics:
A;Gene: yogE

 Query Match 84.8%; Score 28; DB 2; Length 62;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WLIQL 5
|||||
Db 35 WLIQL 39

Search completed: October 14, 2003, 14:38:01
Job time : 9.875 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 14:33:18 ; Search time 4.25 Seconds
(without alignments)
66.391 Million cell updates/sec

Title: US-09-881-490-154
Perfect score: 33
Sequence: 1 WLIQLK 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	29	87.9	207	1	RIB2_BUCAI	P57534 buchnera ap
2	29	87.9	341	1	YB57_METJA	Q58558 methanococc
3	29	87.9	388	1	YBL2_SFV3L	P27403 simian foam
4	29	87.9	620	1	YJB9_YEAST	P47069 saccharomyc
5	29	87.9	1160	1	DP3A_SALTY	P14567 salmonella
6	28	84.8	162	1	GCSH_ORYSA	O22535 oryza sativ
7	28	84.8	174	1	NUGC_MESVI	Q9murl mesostigma
8	28	84.8	180	1	GLUC_CAVPO	P05110 cavia porce
9	28	84.8	180	1	GLUC_HUMAN	P01275 homo sapien
10	28	84.8	180	1	GLUC_MESAU	P01273 mesocricetu
11	28	84.8	180	1	GLUC_MOUSE	P55095 mus musculu
12	28	84.8	180	1	GLUC_OCTDE	P22890 octodon deg
13	28	84.8	180	1	GLUC_RAT	P06883 rattus norv
14	28	84.8	194	1	PUR1_LACCA	P35853 lactobacill
15	28	84.8	209	1	RYL2_YARLI	P41925 yarrowia li
16	28	84.8	226	1	ATP6_CANFA	Q9zz62 canis famil
17	28	84.8	226	1	ATP6_HALGR	P38591 halichoerus
18	28	84.8	226	1	ATP6_HIPAM	Q9zzy6 hippopotamu
19	28	84.8	226	1	ATP6_HYLLA	Q95706 hylobates l
20	28	84.8	226	1	ATP6_PHOVI	Q00521 phoca vitul
21	28	84.8	226	1	ATP6_PIG	Q35915 sus scrofa
22	28	84.8	257	1	NUDC_ECCL6	Q8fb75 escherichia
23	28	84.8	371	1	YB2X_HAEIN	O86233 haemophilus
24	28	84.8	379	1	Y829_BRUME	Q8yhh1 brucella me
25	28	84.8	431	1	RAM1_YEAST	P22007 saccharomyc
26	28	84.8	434	1	CEMA_MARPO	P12211 marchantia
27	28	84.8	482	1	BPI_BOVIN	P17453 bos taurus
28	28	84.8	483	1	BPI_HUMAN	P17213 homo sapien
29	28	84.8	547	1	YOJI_ECOLI	P33941 escherichia
30	28	84.8	576	1	DSBD_PASMU	Q9cp40 pasteurella
31	28	84.8	672	1	YCDR_ECOLI	P75906 escherichia
32	28	84.8	910	1	SYV_SYNY3	Q55522 synechocyst
33	28	84.8	1105	1	YEGE_ECOLI	P38097 escherichia
34	28	84.8	1742	1	MY5C_HUMAN	Q9nqx4 homo sapien
35	28	84.8	4639	1	DYHC_DROME	P37276 drosophila
36	28	84.8	4725	1	DYHC_DICDI	P34036 dictyosteli
37	27	81.8	95	1	NCCY_ALCXX	Q44581 alcaligenes
38	27	81.8	127	1	CRCB_YERPE	Q8zdh2 yersinia pe
39	27	81.8	212	1	IFE4_CAEEL	Q22888 caenorhabdi
40	27	81.8	226	1	ATP6_DUGDU	Q8w9n1 dugong dugo
41	27	81.8	226	1	ATP6_EQUAS	P92480 equus asinu
42	27	81.8	226	1	ATP6_FELCA	P48894 felis silve
43	27	81.8	226	1	ATP6_HORSE	P48662 equus cabal
44	27	81.8	311	1	HFLK_BORBU	O51221 borrelia bu
45	27	81.8	340	1	CELF_VZVD	P09261 varicella-z

ALIGNMENTS

RESULT 1
RIB2_BUCAI
ID RIB2_BUCAI STANDARD; PRT; 207 AA.
AC P57534;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 5-amino-6-(5-phosphoribosylamino)uracil reductase (EC 1.1.1.193) (HTP
DE reductase).
GN RIBD2 OR BU462.
OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
OS symbiotic bacterium).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
OX NCBI_TaxID=118099;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tokyo 1998;
RX MEDLINE=20445173; PubMed=10993077;
RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RT "Genome sequence of the endocellular bacterial symbiont of aphids
RT Buchnera sp. APS.";
RL Nature 407:81-86(2000).
CC -!- CATALYTIC ACTIVITY: 5-amino-6-(5-phosphoribitylamino)uracil +
CC NADP(+) = 5-amino-6-(5-phosphoribosylamino)uracil + NADPH.
CC -!- PATHWAY: Riboflavin biosynthesis; third step.
CC -!- SIMILARITY: BELONGS TO THE HTP REDUCTASE FAMILY.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to licensc@isb-sib.ch).
CC ---
DR EMBL; AP001119; BAB13159.1; -.
DR InterPro; IPR002734; RibD_C.
DR Pfam; PF01872; RibD_C; 1.
DR TIGRFAMs; TIGR00227; ribD_Cterm; 1.
KW Riboflavin biosynthesis; Oxidoreductase; NADP; Complete proteome.
SQ SEQUENCE 207 AA; 23942 MW; DAC973284089D76A CRC64;

Query Match 87.9%; Score 29; DB 1; Length 207;
Best Local Similarity 83.3%; Pred. No. 30;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 WLIQLK 6
              |||:||
Db         90 WLIRLK 95
```

RESULT 2
YB57_METJA
ID YB57_METJA STANDARD; PRT; 341 AA.
AC Q58558;

```
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical protein MJ1157.
GN   MJ1157.
OS   Methanococcus jannaschii.
OC   Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC   Methanocaldococcaceae; Methanocaldococcus.
OX   NCBI_TaxID=2190;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX   MEDLINE=96337999; PubMed=8688087;
RA   Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA   Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA   Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA   Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA   Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA   Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA   Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA   Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT   "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT   jannaschii.";
RL   Science 273:1058-1073(1996).
CC   -!- SIMILARITY: BELONGS TO THE UPF0021 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U67558; AAB99161.1; -.
DR   PIR; E64444; E64444.
DR   TIGR; MJ1157; -.
DR   InterPro; IPR000541; UPF0021.
DR   Pfam; PF01171; ATP_bind3; 1.
DR   TIGRFAMs; TIGR00269; TIGR00269; 1.
DR   PROSITE; PS01263; UPF0021; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   341 AA;  40094 MW;  B87D2AC28D6C18C3 CRC64;

  Query Match             87.9%;  Score 29;  DB 1;  Length 341;
  Best Local Similarity   83.3%;  Pred. No. 50;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              |||:||
Db        333 WLIKLK 338


RESULT 3
YBL2_SFV3L
ID   YBL2_SFV3L     STANDARD;      PRT;   388 AA.
AC   P27403; Q88191;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   BEL-2 protein.
GN   BEL-2.
OS   Simian foamy virus (type 3 / strain LK3) (SFV-3).
OC   Viruses; Retroid viruses; Retroviridae; Spumavirus.
OX   NCBI_TaxID=11644;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92124734; PubMed=1310187;
RA   Renne R., Friedl E., Schweizer M., Fleps U., Turek R.,
RA   Neumann-Haefelin D.;
RT   "Genomic organization and expression of simian foamy virus type 3
RT   (SFV-3).";
RL   Virology 186:597-608(1992).
CC   -!- FUNCTION: MIGHT BE IMPORTANT IN VIRAL LIFE CYCLE.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M74895; AAA47800.1; -.
DR   EMBL; M74895; AAA47797.1; ALT_INIT.
DR   InterPro; IPR004956; Foamy_BEL.
DR   Pfam; PF03274; Foamy_BEL; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   388 AA;  44527 MW;  45D6F74F3FD53966 CRC64;

  Query Match             87.9%;  Score 29;  DB 1;  Length 388;
  Best Local Similarity   83.3%;  Pred. No. 56;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              |||:||
Db        182 WLIRLK 187


RESULT 4
YJB9_YEAST
ID   YJB9_YEAST     STANDARD;      PRT;   620 AA.
AC   P47069;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   Hypothetical 72.4 kDa protein in PET130-CCT3 intergenic region.
GN   YJL019W OR J1310.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   To Van D., Perea J., Jacq C.;
RL   Submitted (SEP-1995) to the EMBL/GenBank/DDBJ databases.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z49294; CAA89309.1; -.
DR   PIR; S56790; S56790.
DR   SGD; S0003556; YJL019W.
DR   GO; GO:0007103; P:spindle pole body duplication (sensu Saccha. . .; IMP.
KW   Hypothetical protein; Transmembrane; ATP-binding.
FT   DOMAIN       88     93       POLY-ASP.
FT   TRANSMEM    155    175       POTENTIAL.
FT   NP_BIND     187    194       ATP (POTENTIAL).
FT   DOMAIN      391    401       POLY-GLN.
SQ   SEQUENCE   620 AA;  72385 MW;  18EC13469B9A5C6B CRC64;

  Query Match             87.9%;  Score 29;  DB 1;  Length 620;
  Best Local Similarity   83.3%;  Pred. No. 90;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              | ||||
Db        579 WFIQLK 584


RESULT 5
```

```
DP3A_SALTY
ID   DP3A_SALTY     STANDARD;      PRT;  1160 AA.
AC   P14567;
DT   01-JAN-1990 (Rel. 13, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   DNA polymerase III alpha subunit (EC 2.7.7.7).
GN   DNAE OR POLC OR STM0231.
OS   Salmonella typhimurium.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Salmonella.
OX   NCBI_TaxID=602;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90008797; PubMed=2676978;
RA   Lancy E.D., Lifsics M.R., Munson P., Maurer R.;
RT   "Nucleotide sequences of dnaE, the gene for the polymerase subunit of
RT   DNA polymerase III in Salmonella typhimurium, and a variant that
RT   facilitates growth in the absence of another polymerase subunit.";
RL   J. Bacteriol. 171:5581-5586(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX   MEDLINE=21534948; PubMed=11677609;
RA   McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA   Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA   Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA   Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA   Waterston R., Wilson R.K.;
RT   "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT   LT2.";
RL   Nature 413:852-856(2001).
CC   -!- FUNCTION: DNA POLYMERASE III IS A COMPLEX, MULTICHAIN ENZYME
CC       RESPONSIBLE FOR MOST OF THE REPLICATIVE SYNTHESIS IN BACTERIA.
CC       THIS DNA POLYMERASE ALSO EXHIBITS 3' TO 5' EXONUCLEASE ACTIVITY.
CC       THE ALPHA CHAIN IS THE DNA POLYMERASE.
CC   -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC       + (DNA)(N).
CC   -!- SUBUNIT: DNA polymerase III contains a core (composed of alpha,
CC       epsilon and theta chains) that associates with a tau subunit. This
CC       core dimerizes to form the POLIII' complex. PolIII' associates
CC       with the gamma complex (composed of gamma, delta, delta', psi and
CC       chi chains) and with the beta chain to form the complete DNA
CC       polymerase III complex. The final composition of the complex
CC       is: (alpha,epsilon,theta)[2]-tau[2]-(gamma,delta,delta',
CC       psi,chi)[2]-beta[4].
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-C FAMILY. DNAE
CC       SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M29701; AAA27057.1; -.
DR   EMBL; M26046; AAA27191.1; -.
DR   EMBL; AE008705; AAL19195.1; -.
DR   StyGene; SG10087; dnaE.
DR   InterPro; IPR004013; PHP_C.
DR   InterPro; IPR003141; PHP_N.
DR   InterPro; IPR004805; PolC_alpha.
DR   InterPro; IPR004365; tRNA_anti.
DR   Pfam; PF02811; PHP_C; 1.
DR   Pfam; PF02231; PHP_N; 1.
DR   Pfam; PF01336; tRNA_anti; 1.
DR   SMART; SM00481; POLIIIAc; 1.
DR   TIGRFAMs; TIGR00594; polc; 1.
KW   Transferase; DNA-directed DNA polymerase; DNA replication;
KW   Complete proteome.
FT   CONFLICT    405    405       D -> S (IN REF. 1; AAA27191).
FT   CONFLICT    782    782       F -> I (IN REF. 1; AAA27191).
FT   CONFLICT    997    997       A -> S (IN REF. 1; AAA27191).
FT   CONFLICT   1127   1127       R -> G (IN REF. 1).
SQ   SEQUENCE   1160 AA;  130217 MW;  2E9651A7C7429533 CRC64;

  Query Match             87.9%;  Score 29;  DB 1;  Length 1160;
  Best Local Similarity   66.7%;  Pred. No. 1.7e+02;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              ||::||
Db        120 WLVELK 125


RESULT 6
GCSH_ORYSA
ID   GCSH_ORYSA     STANDARD;      PRT;   162 AA.
AC   O22535;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Glycine cleavage system H protein, mitochondrial precursor.
GN   GDCSH.
OS   Oryza sativa (Rice).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Milyang 23; TISSUE=Seed;
RA   Lee M.C., Kim C.S., Eun M.Y.;
RT   "Isolation and characterization of H protein from rice.";
RL   Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: THE GLYCINE CLEAVAGE SYSTEM CATALYZES THE DEGRADATION OF
CC       GLYCINE. THE H PROTEIN SHUTTLES THE METHYLAMINE GROUP OF GLYCINE
CC       FROM THE P PROTEIN TO THE T PROTEIN (BY SIMILARITY).
CC   -!- COFACTOR: THE H CHAIN CONTAINS A COVALENTLY-BOUND LIPOYL COFACTOR
CC       (BY SIMILARITY).
CC   -!- SUBUNIT: The glycine cleavage system is composed of four proteins:
CC       P, T, L, and H (By similarity).
CC   -!- SUBCELLULAR LOCATION: Mitochondrial (By similarity).
CC   -!- SIMILARITY: Contains 1 lipoyl-binding domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF022731; AAB82134.1; -.
DR   PIR; T02072; T02072.
DR   HSSP; P16048; 1HTP.
DR   Gramene; O22535; -.
DR   InterPro; IPR002930; GCV_H.
DR   InterPro; IPR003016; Lipoyl.
DR   Pfam; PF01597; GCV_H; 1.
DR   TIGRFAMs; TIGR00527; gcvH; 1.
DR   PROSITE; PS00189; LIPOYL; 1.
KW   Mitochondrion; Transit peptide; Lipoyl.
FT   TRANSIT       1     34       MITOCHONDRION (BY SIMILARITY).
FT   CHAIN        35    162       GLYCINE CLEAVAGE SYSTEM H PROTEIN.
FT   BINDING      97     97       LIPOYL (BY SIMILARITY).
SQ   SEQUENCE   162 AA;  17218 MW;  E212A6AE916EC0FF CRC64;

  Query Match             84.8%;  Score 28;  DB 1;  Length 162;
  Best Local Similarity   66.7%;  Pred. No. 39;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
```

```
              |:|||:|
Db        133 WMIQVK 138


RESULT 7
NUGC_MESVI
ID   NUGC_MESVI     STANDARD;      PRT;   174 AA.
AC   Q9MUR1;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   NAD(P)H-quinone oxidoreductase chain J, chloroplast (EC 1.6.5.-)
DE   (NAD(P)H dehydrogenase, chain J) (NADH-plastoquinone oxidoreductase
DE   subunit J).
GN   NDHJ.
OS   Mesostigma viride.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Mesostigmatophyceae;
OC   Mesostigmatales; Mesostigmataceae; Mesostigma.
OX   NCBI_TaxID=41882;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NIES-296;
RX   MEDLINE=20150907; PubMed=10688199;
RA   Lemieux C., Otis C., Turmel M.;
RT   "Ancestral chloroplast genome in Mesostigma viride reveals an early
RT   branch of green plant evolution.";
RL   Nature 403:649-652(2000).
CC   -!- CATALYTIC ACTIVITY: NAD(P)H + plastoquinone = NAD(P)(+) +
CC       plastoquinol.
CC   -!- SIMILARITY: BELONGS TO THE COMPLEX I 30 kDa SUBUNIT FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF166114; AAF43839.1; -.
DR   InterPro; IPR001268; Complex1_30K.
DR   Pfam; PF00329; complex1_30Kd; 1.
DR   ProDom; PD001581; Complex1_30K; 1.
DR   PROSITE; PS00542; COMPLEX1_30K; 1.
KW   Oxidoreductase; NAD; NADP; Quinone; Plastoquinone; Chloroplast.
SQ   SEQUENCE   174 AA;  20329 MW;  122ABFA9188D47B7 CRC64;

  Query Match             84.8%;  Score 28;  DB 1;  Length 174;
  Best Local Similarity   66.7%;  Pred. No. 42;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              |||::|
Db         24 WLIEMK 29


RESULT 8
GLUC_CAVPO
ID   GLUC_CAVPO     STANDARD;      PRT;   180 AA.
AC   P05110;
DT   13-AUG-1987 (Rel. 05, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Glucagon precursor [Contains: Glicentin-related polypeptide (GRPP);
DE   Glucagon; Glucagon-37 (Oxyntomodulin); Glucagon-like peptide 1 (GLP1);
DE   Glucagon-like peptide 2 (GLP2)].
GN   GCG.
OS   Cavia porcellus (Guinea pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX   NCBI_TaxID=10141;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86248118; PubMed=3755107;
RA   Seino S., Welsh M., Bell G.I., Chan S.J., Steiner D.F.;
RT   "Mutations in the guinea pig preproglucagon gene are restricted to a
RT   specific portion of the prohormone sequence.";
RL   FEBS Lett. 203:25-30(1986).
RN   [2]
RP   SEQUENCE OF 53-81.
RX   MEDLINE=86165412; PubMed=3956884;
RA   Huang C.G., Eng J., Pan Y.-C.E., Hulmes J.D., Yalow R.S.;
RT   "Guinea pig glucagon differs from other mammalian glucagons.";
RL   Diabetes 35:508-512(1986).
RN   [3]
RP   PARTIAL SEQUENCE OF 53-89.
RX   MEDLINE=86017849; PubMed=4048553;
RA   Conlon J.M., Hansen H.F., Schwartz T.W.;
RT   "Primary structure of glucagon and a partial sequence of
RT   oxyntomodulin (glucagon-37) from the guinea pig.";
RL   Regul. Pept. 11:309-320(1985).
CC   -!- FUNCTION: GLUCAGON PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND
CC       RAISES THE BLOOD SUGAR LEVEL.
CC   -!- FUNCTION: GLP2 STIMULATES INTESTINAL GROWTH AND UPREGULATES VILLUS
CC       HEIGHT IN THE SMALL INTESTINE, CONCOMITANT WITH INCREASED CRYPT
CC       CELL PROLIFERATION AND DECREASED ENTEROCYTE APOPTOSIS.
CC   -!- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS
CC       IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.
CC   -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D00014; BAA00010.1; -.
DR   PIR; A24856; GCGP.
DR   HSSP; P01274; 1GCN.
DR   InterPro; IPR000532; Glucagon.
DR   Pfam; PF00123; hormone2; 3.
DR   PRINTS; PR00275; GLUCAGON.
DR   SMART; SM00070; GLUCA; 3.
DR   PROSITE; PS00260; GLUCAGON; 4.
KW   Glucagon family; Hormone; Cleavage on pair of basic residues; Signal.
FT   SIGNAL        1     20
FT   PEPTIDE      21     50       GLICENTIN-RELATED POLYPEPTIDE.
FT   PEPTIDE      53     81       GLUCAGON.
FT   PEPTIDE      53     89       GLUCAGON-37.
FT   PEPTIDE      92    128       GLUCAGON-LIKE PEPTIDE 1.
FT   PROPEP      131    143
FT   PEPTIDE     146    178       GLUCAGON-LIKE PEPTIDE 2.
SQ   SEQUENCE   180 AA;  20972 MW;  702FB181161D2776 CRC64;

  Query Match             84.8%;  Score 28;  DB 1;  Length 180;
  Best Local Similarity   83.3%;  Pred. No. 44;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              |||| |
Db        170 WLIQTK 175


RESULT 9
GLUC_HUMAN
ID   GLUC_HUMAN     STANDARD;      PRT;   180 AA.
AC   P01275;
DT   21-JUL-1986 (Rel. 01, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Glucagon precursor [Contains: Glicentin-related polypeptide (GRPP);
DE   Glucagon; Glucagon-like peptide 1 (GLP1); Glucagon-like peptide 2
DE   (GLP2)].
```

```
GN   GCG.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88330860; PubMed=2901414;
RA   Drucker D.J., Asa S.;
RT   "Glucagon gene expression in vertebrate brain.";
RL   J. Biol. Chem. 263:13475-13478(1988).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86259053; PubMed=3725587;
RA   White J.W., Saunders G.F.;
RT   "Structure of the human glucagon gene.";
RL   Nucleic Acids Res. 14:4719-4730(1986).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=83271477; PubMed=6877358;
RA   Bell G.I., Sanchez-Pescador R., Laybourn P.J., Najarian R.C.;
RT   "Exon duplication and divergence in the human preproglucagon gene.";
RL   Nature 304:368-371(1983).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Pancreas;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [5]
RP   SEQUENCE OF 53-81.
RA   Thomsen J., Kristiansen K., Brunfeldt K., Sundby F.;
RT   "The amino acid sequence of human glucagon.";
RL   FEBS Lett. 21:315-319(1972).
RN   [6]
RP   SEQUENCE OF 98-127.
RX   MEDLINE=89327238; PubMed=2753890;
RA   Orskov C., Bersani M., Johnsen A.H., Hoejrup P., Holst J.J.;
RT   "Complete sequences of glucagon-like peptide-1 from human and pig
RT   small intestine.";
RL   J. Biol. Chem. 264:12826-12829(1989).
RN   [7]
RP   X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS) OF 53-81.
RX   MEDLINE=98334683; PubMed=9667960;
RA   Sturm N.S., Lin Y., Burley S.K., Krstenansky J.L., Ahn J.M.,
RA   Azizeh B.Y., Trivedi D., Hruby V.J.;
RT   "Structure-function studies on positions 17, 18, and 21 replacement
RT   analogues of glucagon: the importance of charged residues and salt
RT   bridges in glucagon biological activity.";
RL   J. Med. Chem. 41:2693-2700(1998).
CC   -!- FUNCTION: GLUCAGON PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND
CC       RAISES THE BLOOD SUGAR LEVEL.
CC   -!- FUNCTION: GLP2 STIMULATES INTESTINAL GROWTH AND UPREGULATES VILLUS
CC       HEIGHT IN THE SMALL INTESTINE, CONCOMITANT WITH INCREASED CRYPT
CC       CELL PROLIFERATION AND DECREASED ENTEROCYTE APOPTOSIS.
CC   -!- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS
CC       IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.
CC   -!- PHARMACEUTICAL: Available under the names Glucagon (Eli Lilly) and
CC       GlucaGen or Glucagon Novo Nordisk (Novo Nordisk). Used to treat
CC       severe hypoglycemia in insulin-dependent diabetics.
CC   -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
CC   -!- DATABASE: NAME=Glucagon at Eli Lilly;
CC       NOTE=Clinical information on Eli Lilly glucagon products;
CC       WWW="http://www.lillyDiabetes.com/Products/PatientInfo.cfm".
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J04040; AAA52567.1; -.
DR   EMBL; X03991; CAA27627.1; -.
DR   EMBL; V01515; CAA24759.1; -.
DR   EMBL; BC005278; AAH05278.1; -.
DR   PIR; A24377; GCHU.
DR   PDB; 1BH0; 18-NOV-98.
DR   PDB; 1D0R; 23-OCT-02.
DR   Genew; HGNC:4191; GCG.
DR   MIM; 138030; -.
DR   MIM; 231530; -.
DR   GO; GO:0005625; C:soluble fraction; TAS.
DR   GO; GO:0008283; P:cell proliferation; TAS.
DR   GO; GO:0007631; P:feeding behavior; TAS.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; TAS.
DR   GO; GO:0007165; P:signal transduction; TAS.
DR   InterPro; IPR000532; Glucagon.
DR   Pfam; PF00123; hormone2; 3.
DR   SMART; SM00070; GLUCA; 3.
DR   PROSITE; PS00260; GLUCAGON; 4.
KW   Glucagon family; Hormone; Cleavage on pair of basic residues; Signal;
KW   Pharmaceutical; 3D-structure; Polymorphism.
FT   SIGNAL       1     20
FT   PEPTIDE     21     50       GLICENTIN-RELATED POLYPEPTIDE.
FT   PEPTIDE     53     81       GLUCAGON.
FT   PROPEP      84     96
FT   PEPTIDE     98    127       GLUCAGON-LIKE PEPTIDE 1.
FT   PROPEP     131    143
FT   PEPTIDE    146    178       GLUCAGON-LIKE PEPTIDE 2.
FT   VARIANT    115    115       A -> V (IN dbSNP:5650).
FT                               /FTId=VAR_014596.
FT   CONFLICT    82     82       K -> N (IN REF. 3).
FT   TURN        59     62
FT   HELIX       63     77
FT   TURN        78     79
SQ   SEQUENCE   180 AA;  20909 MW;  7A99EEC629B2862C CRC64;

  Query Match             84.8%;  Score 28;  DB 1;  Length 180;
  Best Local Similarity   83.3%;  Pred. No. 44;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              |||| |
Db        170 WLIQTK 175


RESULT 10
GLUC_MESAU
ID   GLUC_MESAU     STANDARD;      PRT;   180 AA.
AC   P01273;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Glucagon precursor [Contains: Glicentin-related polypeptide (GRPP);
DE   Glucagon; Glucagon-like peptide 1 (GLP1); Glucagon-like peptide 2
```

```
DE   (GLP2)].
GN   GCG.
OS   Mesocricetus auratus (Golden hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Mesocricetus.
OX   NCBI_TaxID=10036;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=83167563; PubMed=6835407;
RA   Bell G.I., Santerre R.F., Mullenbach G.T.;
RT   "Hamster preproglucagon contains the sequence of glucagon and two
RT   related peptides.";
RL   Nature 302:716-718(1983).
RN   [2]
RP   REVISIONS TO 12-15.
RA   Bell G.I.;
RL   Submitted (XXX-1985) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: GLUCAGON PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND
CC       RAISES THE BLOOD SUGAR LEVEL.
CC   -!- FUNCTION: GLP2 STIMULATES INTESTINAL GROWTH AND UPREGULATES VILLUS
CC       HEIGHT IN THE SMALL INTESTINE, CONCOMITANT WITH INCREASED CRYPT
CC       CELL PROLIFERATION AND DECREASED ENTEROCYTE APOPTOSIS.
CC   -!- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS
CC       IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.
CC   -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J00059; AAA37074.1; -.
DR   HSSP; P01274; 1GCN.
DR   InterPro; IPR000532; Glucagon.
DR   Pfam; PF00123; hormone2; 3.
DR   SMART; SM00070; GLUCA; 3.
DR   PROSITE; PS00260; GLUCAGON; 4.
KW   Glucagon family; Hormone; Cleavage on pair of basic residues; Signal.
FT   SIGNAL       1     20
FT   PEPTIDE      21    50       GLICENTIN-RELATED POLYPEPTIDE.
FT   PEPTIDE      53    81       GLUCAGON.
FT   PROPEP       84    89
FT   PEPTIDE      92    128      GLUCAGON-LIKE PEPTIDE 1.
FT   PROPEP       131   143
FT   PEPTIDE      146   178      GLUCAGON-LIKE PEPTIDE 2.
SQ   SEQUENCE   180 AA;  20954 MW;  02791B49D7AADD4B CRC64;

  Query Match             84.8%;  Score 28;  DB 1;  Length 180;
  Best Local Similarity   83.3%;  Pred. No. 44;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              |||| |
Db        170 WLIQTK 175


RESULT 11
GLUC_MOUSE
ID   GLUC_MOUSE     STANDARD;      PRT;   180 AA.
AC   P55095;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Glucagon precursor [Contains: Glicentin-related polypeptide (GRPP);
DE   Glucagon; Glucagon-like peptide 1 (GLP1); Glucagon-like peptide 2
DE   (GLP2)].
GN   GCG.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Pancreatic islets;
RX   MEDLINE=95247722; PubMed=7730317;
RA   Rothenberg M.E., Eilertson C.D., Klein K., Zhou Y., Linberg I.,
RA   McDonald J.K., Mackin R.B., Noe B.D.;
RT   "Processing of mouse proglucagon by recombinant prohormone convertase
RT   1 and immunopurified prohormone convertase 2 in vitro.";
RL   J. Biol. Chem. 270:10136-10146(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Shamsadin R., Knepel W.;
RT   "Mouse glucagon full length cDNA.";
RL   Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: GLUCAGON PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND
CC       RAISES THE BLOOD SUGAR LEVEL.
CC   -!- FUNCTION: GLP2 STIMULATES INTESTINAL GROWTH AND UPREGULATES VILLUS
CC       HEIGHT IN THE SMALL INTESTINE, CONCOMITANT WITH INCREASED CRYPT
CC       CELL PROLIFERATION AND DECREASED ENTEROCYTE APOPTOSIS.
CC   -!- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS
CC       IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.
CC   -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z46845; CAA86902.1; -.
DR   EMBL; AF276754; AAK96898.1; -.
DR   PIR; A57294; A57294.
DR   HSSP; P01274; 1GCN.
DR   MGD; MGI:95674; Gcg.
DR   InterPro; IPR000532; Glucagon.
DR   Pfam; PF00123; hormone2; 3.
DR   PRINTS; PR00275; GLUCAGON.
DR   SMART; SM00070; GLUCA; 3.
DR   PROSITE; PS00260; GLUCAGON; 4.
KW   Glucagon family; Hormone; Cleavage on pair of basic residues; Signal.
FT   SIGNAL       1     20       BY SIMILARITY.
FT   PEPTIDE      21    50       GLICENTIN-RELATED POLYPEPTIDE.
FT   PEPTIDE      53    81       GLUCAGON.
FT   PROPEP       84    89
FT   PEPTIDE      92    128      GLUCAGON-LIKE PEPTIDE 1.
FT   PROPEP       131   143
FT   PEPTIDE      146   178      GLUCAGON-LIKE PEPTIDE 2.
SQ   SEQUENCE   180 AA;  20906 MW;  595AA6DD9A589950 CRC64;

  Query Match             84.8%;  Score 28;  DB 1;  Length 180;
  Best Local Similarity   83.3%;  Pred. No. 44;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              |||| |
Db        170 WLIQTK 175


RESULT 12
GLUC_OCTDE
ID   GLUC_OCTDE     STANDARD;      PRT;   180 AA.
AC   P22890;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Glucagon precursor [Contains: Glicentin-related polypeptide (GRPP);
DE   Glucagon; Glucagon-like peptide 1 (GLP1); Glucagon-like peptide 2
DE   (GLP2)].
GN   GCG.
```

```
OS   Octodon degus (Degu).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Hystricognathi; Octodontidae; Octodon.
OX   NCBI_TaxID=10160;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91155952; PubMed=2293024;
RA   Nishi M., Steiner D.F.;
RT   "Cloning of complementary DNAs encoding islet amyloid polypeptide,
RT   insulin, and glucagon precursors from a New World rodent, the degu.
RT   Octodon degus.";
RL   Mol. Endocrinol. 4:1192-1198(1990).
CC   -!- FUNCTION: GLUCAGON PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND
CC       RAISES THE BLOOD SUGAR LEVEL.
CC   -!- FUNCTION: GLP2 STIMULATES INTESTINAL GROWTH AND UPREGULATES VILLUS
CC       HEIGHT IN THE SMALL INTESTINE, CONCOMITANT WITH INCREASED CRYPT
CC       CELL PROLIFERATION AND DECREASED ENTEROCYTE APOPTOSIS.
CC   -!- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS
CC       IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.
CC   -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M57688; AAA40588.1; -.
DR   PIR; C36118; GCRTDU.
DR   HSSP; P01274; 1GCN.
DR   InterPro; IPR000532; Glucagon.
DR   Pfam; PF00123; hormone2; 3.
DR   PRINTS; PR00275; GLUCAGON.
DR   SMART; SM00070; GLUCA; 3.
DR   PROSITE; PS00260; GLUCAGON; 4.
KW   Glucagon family; Hormone; Cleavage on pair of basic residues; Signal;
KW   Amidation.
FT   SIGNAL       1     20
FT   PEPTIDE      21    50       GLICENTIN-RELATED POLYPEPTIDE.
FT   PEPTIDE      53    81       GLUCAGON.
FT   PROPEP       84    89
FT   PEPTIDE      92    127      GLUCAGON-LIKE PEPTIDE 1.
FT   PROPEP      131    142
FT   PEPTIDE     146    178      GLUCAGON-LIKE PEPTIDE 2.
FT   MOD_RES     127    127      AMIDATION (G-126 PROVIDE AMIDE GROUP).
SQ   SEQUENCE   180 AA;  21165 MW;  6E8836160A9A3051 CRC64;

  Query Match             84.8%;  Score 28;  DB 1;  Length 180;
  Best Local Similarity   83.3%;  Pred. No. 44;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              |||| |
Db        170 WLIQTK 175


RESULT 13
GLUC_RAT
ID   GLUC_RAT       STANDARD;      PRT;   180 AA.
AC   P06883;
DT   01-JAN-1988 (Rel. 06, Created)
DT   01-JAN-1988 (Rel. 06, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Glucagon precursor [Contains: Glicentin-related polypeptide (GRPP);
DE   Glucagon; Glucagon-like peptide 1 (GLP1); Glucagon-like peptide 2
DE   (GLP2)].
GN   GCG.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85054853; PubMed=6094539;
RA   Heinrich G., Gros P., Habener J.F.;
RT   "Glucagon gene sequence. Four of six exons encode separate functional
RT   domains of rat pre-proglucagon.";
RL   J. Biol. Chem. 259:14082-14087(1984).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85051023; PubMed=6548696;
RA   Heinrich G., Gros P., Lund P.K., Bentley R.C., Habener J.F.;
RT   "Pre-proglucagon messenger ribonucleic acid: nucleotide and encoded
RT   amino acid sequences of the rat pancreatic complementary
RT   deoxyribonucleic acid.";
RL   Endocrinology 115:2176-2181(1984).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86304324; PubMed=3528148;
RA   Mojsov S., Heinrich G., Wilson I.B., Ravazzola M., Orci L.,
RA   Habener J.F.;
RT   "Preproglucagon gene expression in pancreas and intestine diversifies
RT   at the level of post-translational processing.";
RL   J. Biol. Chem. 261:11880-11889(1986).
CC   -!- FUNCTION: GLUCAGON PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND
CC       RAISES THE BLOOD SUGAR LEVEL.
CC   -!- FUNCTION: GLP2 STIMULATES INTESTINAL GROWTH AND UPREGULATES VILLUS
CC       HEIGHT IN THE SMALL INTESTINE, CONCOMITANT WITH INCREASED CRYPT
CC       CELL PROLIFERATION AND DECREASED ENTEROCYTE APOPTOSIS.
CC   -!- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS
CC       IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.
CC   -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; K02813; AAA41235.1; -.
DR   EMBL; K02809; AAA41235.1; JOINED.
DR   EMBL; K02810; AAA41235.1; JOINED.
DR   EMBL; K02811; AAA41235.1; JOINED.
DR   EMBL; K02812; AAA41235.1; JOINED.
DR   PIR; A22655; GCRT.
DR   HSSP; P01274; 1GCN.
DR   InterPro; IPR000532; Glucagon.
DR   Pfam; PF00123; hormone2; 3.
DR   PRINTS; PR00275; GLUCAGON.
DR   SMART; SM00070; GLUCA; 3.
DR   PROSITE; PS00260; GLUCAGON; 4.
KW   Glucagon family; Hormone; Cleavage on pair of basic residues; Signal.
FT   SIGNAL       1     20
FT   PEPTIDE      21    50       GLICENTIN-RELATED POLYPEPTIDE.
FT   PEPTIDE      53    81       GLUCAGON.
FT   PROPEP       84    89
FT   PEPTIDE      92    128      GLUCAGON-LIKE PEPTIDE 1.
FT   PROPEP      131    143
FT   PEPTIDE     146    178      GLUCAGON-LIKE PEPTIDE 2.
SQ   SEQUENCE   180 AA;  20846 MW;  76931409D03C7978 CRC64;

  Query Match             84.8%;  Score 28;  DB 1;  Length 180;
  Best Local Similarity   83.3%;  Pred. No. 44;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              |||| |
Db        170 WLIQTK 175


RESULT 14
PUR1_LACCA
```

```
ID   PURL_LACCA     STANDARD;      PRT;   194 AA.
AC   P35853;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Amidophosphoribosyltransferase precursor (EC 2.4.2.14) (Glutamine
DE   phosphoribosylpyrophosphate amidotransferase) (ATASE) (GPATase)
DE   (Fragment).
GN   PURF.
OS   Lactobacillus casei.
OC   Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC   Lactobacillus.
OX   NCBI_TaxID=1582;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93012962; PubMed=1398079;
RA   Gu Z.-M., Martindale D.W., Lee B.H.;
RT   "Isolation and complete sequence of the purL gene encoding FGAM
RT   synthase II in Lactobacillus casei.";
RL   Gene 119:123-126(1992).
RN   [2]
RP   ERRATUM.
RX   MEDLINE=94040790; PubMed=8224889;
RA   Gu Z.-M., Martindale D.W., Lee B.H.;
RL   Gene 133:147-147(1993).
CC   -!- CATALYTIC ACTIVITY: 5-phospho-beta-D-ribosylamine + diphosphate +
CC       L-glutamate = L-glutamine + 5-phospho-alpha-D-ribose 1-diphosphate
CC       + H(2)O.
CC   -!- PATHWAY: De novo purine biosynthesis; first step.
CC   -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE
CC       PURINE/PYRIMIDINE PHOSPHORIBOSYLTRANSFERASE FAMILY.
CC   -!- SIMILARITY: Contains 1 type-2 glutamine amidotransferase domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M85265; AAC36948.1; -.
DR   PIR; PC1136; PC1136.
DR   HSSP; P00497; 1AO0.
DR   MEROPS; C44.001; -.
DR   InterPro; IPR000583; GATase_2.
DR   InterPro; IPR002375; Pr/py_rp_transf.
DR   Pfam; PF00310; GATase_2; 1.
DR   PROSITE; PS00103; PUR_PYR_PR_TRANSFER; PARTIAL.
DR   PROSITE; PS00443; GATASE_TYPE_II; 1.
KW   Purine biosynthesis; Transferase; Glycosyltransferase.
FT   PROPEP        1     11       BY SIMILARITY.
FT   CHAIN        12   >194       AMIDOPHOSPHORIBOSYLTRANSFERASE.
FT   ACT_SITE     12     12       GATASE (INVOLVED IN SUBSTRATE BINDING AND
FT                                CATALYSIS) (BY SIMILARITY).
FT   NON_TER     194    194
SQ   SEQUENCE   194 AA;  21144 MW;  4A788CBC5365D5EC CRC64;

  Query Match             84.8%;  Score 28;  DB 1;  Length 194;
  Best Local Similarity   83.3%;  Pred. No. 47;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              || |||
Db        154 WLTQLK 159


RESULT 15
RYL2_YARLI
ID   RYL2_YARLI     STANDARD;      PRT;   209 AA.
AC   P41925;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Ras-like GTP-binding protein RYL2.
GN   RYL2.
OS   Yarrowia lipolytica (Candida lipolytica).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Dipodascaceae; Yarrowia.
OX   NCBI_TaxID=4952;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 20460 / W29;
RA   Pertuiset B., Beckerich J.M., Gaillardin C.;
RL   Submitted (XXX-1992) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: PROTEIN TRANSPORT. PROBABLY INVOLVED IN VESICULAR
CC       TRAFFIC (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE SMALL GTPASE SUPERFAMILY. RAB FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L06970; AAA35246.1; -.
DR   HSSP; P36017; 1EK0.
DR   InterPro; IPR003579; GTPase_Rab.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00175; RAB; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
KW   GTP-binding; Lipoprotein; Prenylation; Protein transport.
FT   NP_BIND      12     19       GTP (BY SIMILARITY).
FT   NP_BIND      60     64       GTP (BY SIMILARITY).
FT   NP_BIND     118    121       GTP (BY SIMILARITY).
FT   DOMAIN       34     42       EFFECTOR REGION (PROBABLE).
FT   LIPID       208    208       GERANYL-GERANYL (BY SIMILARITY).
FT   LIPID       209    209       GERANYL-GERANYL (BY SIMILARITY).
SQ   SEQUENCE   209 AA;  23782 MW;  2C4697DCD084844A CRC64;

  Query Match             84.8%;  Score 28;  DB 1;  Length 209;
  Best Local Similarity   66.7%;  Pred. No. 51;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              ||::||
Db         99 WLLELK 104


Search completed: October 14, 2003, 14:34:04
Job time : 5.41667 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 14:33:18 ; Search time 19.875 Seconds
(without alignments)
77.903 Million cell updates/sec

Title: US-09-881-490-154
Perfect score: 33
Sequence: 1 WLIQLK 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	33	100.0	136	16	Q8EI76	Q8ei76 shewanella
2	33	100.0	555	16	Q8G3G5	Q8g3q5 bifidobacte
3	33	100.0	5779	5	Q8IBS0	Q8ibs0 plasmodium
4	32	97.0	116	17	Q96ZA6	Q96za6 sulfolobus
5	32	97.0	182	17	Q97ZU8	Q97zu8 sulfolobus
6	30	90.9	1211	10	Q8GZF0	Q8gzf0 glycine max
7	29	87.9	286	16	Q9CMK7	Q9cmk7 pasteurella
8	29	87.9	287	16	Q98QY4	Q98qy4 mycoplasma
9	29	87.9	369	2	Q93CS6	Q93cs6 shigella bo
10	29	87.9	371	10	Q9M9K8	Q9m9k8 arabidopsis
11	29	87.9	390	3	O42862	O42862 schizosacch
12	29	87.9	415	10	Q9SFF6	Q9sff6 arabidopsis
13	29	87.9	416	10	O80731	O80731 arabidopsis
14	29	87.9	420	10	Q9M1R8	Q9m1r8 arabidopsis
15	29	87.9	422	10	O65250	O65250 arabidopsis
16	29	87.9	433	16	Q8G5Q6	Q8g5q6 bifidobacte
17	29	87.9	462	16	Q8RIM2	Q8rim2 fusobacteri
18	29	87.9	756	10	Q9SHI4	Q9shi4 arabidopsis
19	29	87.9	769	10	Q8RVG2	Q8rvg2 daucus caro
20	29	87.9	819	5	Q9VXJ2	Q9vxj2 drosophila
21	29	87.9	819	5	Q8SXM2	Q8sxm2 drosophila
22	29	87.9	823	5	Q9NHF0	Q9nhf0 drosophila
23	29	87.9	943	13	Q8JGT4	Q8jgt4 xenopus lae
24	29	87.9	1051	11	Q91XT4	Q91xt4 mus musculu
25	29	87.9	1058	11	Q925T1	Q925t1 rattus norv
26	29	87.9	1160	16	Q8Z9A0	Q8z9a0 salmonella
27	29	87.9	1195	10	Q8H6U5	Q8h6u5 zea mays (m
28	29	87.9	1208	10	Q8H6U8	Q8h6u8 zea mays (m
29	29	87.9	1222	10	Q8H6V1	Q8h6v1 zea mays (m
30	29	87.9	1222	10	Q8H6U6	Q8h6u6 zea mays (m
31	29	87.9	1226	10	Q8H6V0	Q8h6v0 zea mays (m
32	29	87.9	1247	10	Q8H6U9	Q8h6u9 zea mays (m
33	29	87.9	1251	10	Q8H6V3	Q8h6v3 zea mays (m
34	29	87.9	1251	10	Q8H6U7	Q8h6u7 zea mays (m
35	29	87.9	1251	10	Q8GS26	Q8gs26 zea mays (m
36	29	87.9	1535	5	Q9VPF0	Q9vpf0 drosophila
37	28	84.8	59	16	Q8YW77	Q8yw77 anabaena sp
38	28	84.8	62	16	Q9CFM6	Q9cfm6 lactococcus
39	28	84.8	65	12	Q8VAR2	Q8var2 white spot
40	28	84.8	66	9	Q38245	Q38245 lactococcus
41	28	84.8	108	8	Q9ZYN6	Q9zyn6 caprimulgus
42	28	84.8	146	16	Q9CHG1	Q9chg1 lactococcus
43	28	84.8	166	8	Q34802	Q34802 hylobates s
44	28	84.8	171	2	Q44628	Q44628 brucella ab
45	28	84.8	177	2	Q8VV80	Q8vv80 colwellia m

ALIGNMENTS

RESULT 1
Q8EI76
ID Q8EI76 PRELIMINARY; PRT; 136 AA.
AC Q8EI76;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN SO0971.
OS Shewanella oneidensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Alteromonadaceae; Shewanella.
OX NCBI_TaxID=70863;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MR-1;
RX MEDLINE=22297686; PubMed=12368813;
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,
RA Meyer T., Tsapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
RA Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
RA Vamathevan J., Weidman J., Impraim M., Lee K., Berry K., Lee C.,
RA Mueller J., Khouri H., Gill J., Utterback T.R., McDonald L.A.,
RA Feldblyum T.V., Smith H.O., Venter J.C., Nealson K.H., Fraser C.M.;
RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
RT Shewanella oneidensis.";
RL Nat. Biotechnol. 20:1118-1123(2002).
DR EMBL; AE015542; AAN54045.1; -.
DR TIGR; SO0971; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 136 AA; 15529 MW; 4207109109796050 CRC64;

Query Match 100.0%; Score 33; DB 16; Length 136;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WLIQLK 6
||||||

```
Db         70 WLIQLK 75


RESULT 2
Q8G3G5
ID   Q8G3G5      PRELIMINARY;      PRT;   555 AA.
AC   Q8G3G5;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical proteasome-associated protein.
GN   BL1795.
OS   Bifidobacterium longum.
OC   Bacteria; Actinobacteria; Actinobacteridae; Bifidobacteriales;
OC   Bifidobacteriaceae; Bifidobacterium.
OX   NCBI_TaxID=216816;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NCC 2705;
RX   MEDLINE=22294977; PubMed=12381787;
RA   Schell M.A., Karmirantzou M., Snel B., Vilanova D., Berger B.,
RA   Pessi G., Zwahlen M.-C., Desiere F., Bork P., Delley M.,
RA   Pridmore R.D., Arigoni F.;
RT   "The genome sequence of Bifidobacterium longum reflects its adaptation
RT   to the human gastrointestinal tract.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:14422-14427(2002).
DR   EMBL; AE014813; AAN25578.1; -.
KW   Hypothetical protein; Proteasome; Complete proteome.
SQ   SEQUENCE   555 AA;  61463 MW;  2B608908FE89986A CRC64;

  Query Match             100.0%;  Score 33;  DB 16;  Length 555;
  Best Local Similarity   100.0%;  Pred. No. 71;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              ||:|||
Db        316 WLIQLK 321


RESULT 3
Q8IBS0
ID   Q8IBS0      PRELIMINARY;      PRT;  5779 AA.
AC   Q8IBS0;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein.
GN   MAL7P1.89.
OS   Plasmodium falciparum (isolate 3D7).
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=36329;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Seeger K., Murphy L., Harris D., Berriman M., Pain A., Hall N.,
RA   Quail M., Barrell B.;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL844506; CAD50928.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   5779 AA;  690899 MW;  F930AC1A8C7B6E4C CRC64;

  Query Match             100.0%;  Score 33;  DB 5;  Length 5779;
  Best Local Similarity   100.0%;  Pred. No. 6.6e-02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              ||||||
Db       1529 WLIQLK 1534


RESULT 4
Q96ZA6
ID   Q96ZA6      PRELIMINARY;      PRT;   116 AA.
AC   Q96ZA6;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Hypothetical protein ST1927.
GN   ST1927.
OS   Sulfolobus tokodaii.
OC   Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC   Sulfolobus.
OX   NCBI_TaxID=111955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JCM 10545 / 7;
RX   PubMed=11572479;
RA   Kawarabayasi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA   Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S.,
RA   Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA   Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA   Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA   Oshima T., Kikuchi H.;
RT   "Complete genome sequence of an aerobic thermoacidophilic
RT   Crenarchaeon, Sulfolobus tokodaii strain7.";
RL   DNA Res. 8:123-140(2001).
DR   EMBL; AP000988; BAB67020.1; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   116 AA;  12932 MW;  097F648F32597C42 CRC64;

  Query Match             97.0%;  Score 32;  DB 17;  Length 116;
  Best Local Similarity   83.3%;  Pred. No. 26;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              ||:|||
Db         29 WLVQLK 34


RESULT 5
Q97ZU8
ID   Q97ZU8      PRELIMINARY;      PRT;   182 AA.
AC   Q97ZU8;
DT   01-OCT-2001 (TrEMBLrel. 18, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE   Hypothetical protein SSO0479.
GN   SSO0479.
OS   Sulfolobus solfataricus.
OC   Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC   Sulfolobus.
OX   NCBI_TaxID=2287;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 35092 / DSM 1617 / P2;
RX   MEDLINE=21332296; PubMed=11427726;
RA   She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA   Awayez M.J., Chan-Weiher C.C.-Y., Clausen I.G., Curtis B.A.,
RA   De Moors A., Erauso G., Fletcher C., Gordon P.M.K.,
RA   Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA   Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA   Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA   Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
RT   "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
DR   EMBL; AE006679; AAK40800.1; -.
DR   InterPro; IPR000051; SAM_bind.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   182 AA;  21277 MW;  E97F7EB93D725BCF CRC64;

  Query Match             97.0%;  Score 32;  DB 17;  Length 182;
  Best Local Similarity   83.3%;  Pred. No. 40;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              ||:|||
Db        110 WLVQLK 115
```

```
RESULT 6
Q8GZF0
ID   Q8GZF0      PRELIMINARY;      PRT:  1211 AA.
AC   Q8GZF0;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Resistance protein KR4.
OS   Glycine max (Soybean).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX   NCBI_TaxID=3847;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Kefeng-1;
RA   Wang B., Zhang J., Chen S.;
RT   "Resistance gene homolog (KR4) from soybean.";
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF502080; AAO15846.1; -.
SQ   SEQUENCE   1211 AA;  136972 MW;  8ABFAD93B655F6A8 CRC64;

  Query Match             90.9%;  Score 30;  DB 10;  Length 1211;
  Best Local Similarity   83.3%;  Pred. No. 6.6e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              |||:||
Db         71 WLIELK 76


RESULT 7
Q9CMK7
ID   Q9CMK7      PRELIMINARY;      PRT:   286 AA.
AC   Q9CMK7;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Hypothetical protein PM0814.
GN   PM0814.
OS   Pasteurella multocida.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Pasteurella.
OX   NCBI_TaxID=747;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Pm70;
RX   MEDLINE=21145866; PubMed=11248100;
RA   May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT   "Complete genomic sequence of Pasteurella multocida Pm70.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
DR   EMBL; AE006120; AAK02898.1; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   286 AA;  32041 MW;  C0E2B335342EDC03 CRC64;

  Query Match             87.9%;  Score 29;  DB 16;  Length 286;
  Best Local Similarity   83.3%;  Pred. No. 2.7e-02;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              || |||
Db        145 WLFQLK 150


RESULT 8
Q98QY4
ID   Q98QY4      PRELIMINARY;      PRT:   287 AA.
AC   Q98QY4;
DT   01-OCT-2001 (TrEMBLrel. 18, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Hypothetical protein MYPU_2260.
GN   MYPU_2260.
OS   Mycoplasma pulmonis.
OC   Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX   NCBI_TaxID=2107;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=UAB CTIP;
RX   MEDLINE=21267165; PubMed=11353084;
RA   Chambaud I., Heilig R., Ferris S., Barbe V., Samson D., Galisson F.,
RA   Moszer I., Dybvig K., Wroblewski H., Viari A., Rocha E.P.C.,
RA   Blanchard A.;
RT   "The complete genome sequence of the murine respiratory pathogen
RT   Mycoplasma pulmonis.";
RL   Nucleic Acids Res. 29:2145-2153(2001).
DR   EMBL; AL445563; CAC13399.1; -.
DR   MypuList; MYPU_2260; -.
DR   InterPro; IPR003802; DUF199.
DR   Pfam; PF02650; DUF199; 1.
DR   TIGRFAMs; TIGR00647; MG103; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   287 AA;  34449 MW;  DA02B90205D4BB4D CRC64;

  Query Match             87.9%;  Score 29;  DB 16;  Length 287;
  Best Local Similarity   83.3%;  Pred. No. 2.8e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              |||:||
Db        272 WLIKLK 277


RESULT 9
Q93CS6
ID   Q93CS6      PRELIMINARY;      PRT:   369 AA.
AC   Q93CS6;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Putative glycosyl transferase.
GN   WBAT.
OS   Shigella boydii.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Shigella.
OX   NCBI_TaxID=621;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21481970; PubMed=11598067;
RA   Wang L., Qu W., Reeves P.R.;
RT   "Sequence Analysis of Four Shigella boydii O-Antigen Loci: Implication
RT   for Escherichia coli and Shigella Relationships.";
RL   Infect. Immun. 69:6923-6930(2001).
DR   EMBL; AF402314; AAL27338.1; -.
DR   InterPro; IPR001296; Glyco_trans_1.
DR   Pfam; PF00534; Glycos_transf_1; 1.
KW   Transferase.
SQ   SEQUENCE   369 AA;  42760 MW;  3E2CCE371A6BA0BB CRC64;

  Query Match             87.9%;  Score 29;  DB 2;  Length 369;
  Best Local Similarity   83.3%;  Pred. No. 3.5e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              |||:||
Db        361 WLIKLK 366


RESULT 10
Q9M9K8
ID   Q9M9K8      PRELIMINARY;      PRT:   371 AA.
AC   Q9M9K8;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
```

```
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Putative pectinacetylesterase.
GN   F10A16.21.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RA   Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B., Wu D.,
RA   Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
RA   Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RT   "Arabidopsis thaliana chromosome III BAC F10A16 genomic sequence.";
RL   Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC012393; AAF26093.1; -.
DR   InterPro; IPR004963; Pec_acetylest.
DR   Pfam; PF03283; PAE; 1.
SQ   SEQUENCE   371 AA;  41277 MW;  7E077A8B56DA8842 CRC64;

  Query Match             87.9%;  Score 29;  DB 10;  Length 371;
  Best Local Similarity   83.3%;  Pred. No. 3.5e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              |||||:
Db         88 WLIQLE 93


RESULT 11
O42862
ID   O42862      PRELIMINARY;      PRT;   390 AA.
AC   O42862;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical 45.3 kDa protein.
GN   SPAC25A8.02.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972h-;
RA   Skelton J., Churcher C.M., Barrell B.G., Rajandream M.A., Wood V.;
RL   Submitted (FEB-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL021809; CAA16952.1; -.
DR   GeneDB_SPombe; SPAC25A8.02; -.
KW   Hypothetical protein.
SQ   SEQUENCE   390 AA;  45304 MW;  203FC1DD42238E09 CRC64;

  Query Match             87.9%;  Score 29;  DB 3;  Length 390;
  Best Local Similarity   66.7%;  Pred. No. 3.7e-02;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              ||:||:
Db        333 WLVQLR 338


RESULT 12
Q9SFF6
ID   Q9SFF6      PRELIMINARY;      PRT;   415 AA.
AC   Q9SFF6;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Putative pectinacetylesterase (AT3g05910/F2O10_3).
GN   F2O10.13 OR AT3G05910.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RA   Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.,
RA   Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
RA   Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RT   "Arabidopsis thaliana chromosome III BAC F2O10 genomic sequence.";
RL   Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA   Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA   Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA   Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA   Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA   Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA   Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA   Ecker J.R.;
RT   "Arabidopsis cDNA clones.";
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Yamada K., Liu S.X., Pham P.K., Banh J., Banno F., Dale J.M.,
RA   Goldsmith A.D., Jiang P.X., Lee J.M., Onodera C.S., Quach H.L.,
RA   Tang C., Toriumi M., Yamamura Y., Yu G., Yu S., Bowser L.,
RA   Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA   Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Koesema E., Lam B.,
RA   Lin J., Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA   Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Tracy S.E.,
RA   Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT   "Full Length cDNA of gene F2O10.13 (GI:6671966).";
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
RA   Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
RA   Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
RA   Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA   Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA   Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA   Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.,
RA   Davis R.W., Ecker J.R., Theologis A.;
RT   "Arabidopsis Open Reading Frame (ORF) Clones.";
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC013454; AAF23225.1; -.
DR   EMBL; AF428303; AAL16135.1; -.
DR   EMBL; AY050847; AAK92782.1; -.
DR   EMBL; AY096751; AAM20385.1; -.
DR   InterPro; IPR004963; Pec_acetylest.
DR   Pfam; PF03283; PAE; 1.
SQ   SEQUENCE   415 AA;  46326 MW;  4131CC9C66D35269 CRC64;

  Query Match             87.9%;  Score 29;  DB 10;  Length 415;
  Best Local Similarity   83.3%;  Pred. No. 3.9e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              |||||:
Db         88 WLIQLE 93


RESULT 13
O80731
ID   O80731      PRELIMINARY;      PRT;   416 AA.
AC   O80731;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Putative pectinesterase (At2g46930/F14M4.24).
```

```
GN   AT2G46930.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=20083487; PubMed=10617197;
RA   Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA   Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA   Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA   Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA   Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA   Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA   Salzberg S.L., Fraser C.M., Venter J.C.;
RT   "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 402:761-768(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RA   Lin X.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA   Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA   Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA   Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA   Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA   Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA   Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Rheologis A.,
RA   Ecker J.R.;
RT   "Arabidopsis cDNA clones.";
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC004411; AAC34238.1; -.
DR   EMBL; AY052671; AAK96575.1; -.
DR   InterPro; IPR004963; Pec_acetylest.
DR   InterPro; IPR002016; Peroxidase.
DR   Pfam; PF03283; PAE; 1.
DR   PROSITE; PS00435; PEROXIDASE_1; 1.
SQ   SEQUENCE   416 AA;  45944 MW;  20D4EE2B75EEB387 CRC64;

  Query Match             87.9%;  Score 29;  DB 10;  Length 416;
  Best Local Similarity   83.3%;  Pred. No. 3.9e-02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              |||||:
Db         89 WLIQLE 94


RESULT 14
Q9M1R8
ID   Q9M1R8      PRELIMINARY;      PRT;   420 AA.
AC   Q9M1R8;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Pectinacetylesterase precursor-like protein.
GN   T17J13.20.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Rieger M., Mueller-Auer S., Zipp M., Schaefer M., Mewes H.W.,
RA   Lemcke K., Mayer K.F.X., Quetier F., Salanoubat M.;
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL138651; CAB71866.1; -.
DR   InterPro; IPR004963; Pec_acetylest.
DR   Pfam; PF03283; PAE; 1.
SQ   SEQUENCE   420 AA;  46399 MW;  42E6E8AF8A5F700A CRC64;

  Query Match             87.9%;  Score 29;  DB 10;  Length 420;
  Best Local Similarity   83.3%;  Pred. No. 4e-02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              |||||:
Db         91 WLIQLE 96


RESULT 15
O65250
ID   O65250      PRELIMINARY;      PRT;   422 AA.
AC   O65250;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   F21E10.11 protein.
GN   F21E10.11.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RA   Davidson S., Rohlfing T., David M., O'Brian D.;
RT   "The sequence of A. thaliana F21E10.";
RL   Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RA   WashU;
RT   "The A. thaliana Genome Sequencing Project.";
RL   Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RA   Waterston R.;
RL   Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RA   Wilson R.;
RL   Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF058914; AAC13595.1; -.
DR   InterPro; IPR001064; Crystallin.
DR   InterPro; IPR004963; Pec_acetylest.
DR   Pfam; PF03283; PAE; 1.
DR   PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; 1.
SQ   SEQUENCE   422 AA;  47340 MW;  D9786B080A439BE4 CRC64;

  Query Match             87.9%;  Score 29;  DB 10;  Length 422;
  Best Local Similarity   83.3%;  Pred. No. 4e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 WLIQLK 6
              |||||:
Db         89 WLIQLE 94


Search completed: October 14, 2003, 14:36:52
Job time : 21.875 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 14:33:19 ; Search time 38.625 Seconds
(without alignments)
36.985 Million cell updates/sec

Title: US-09-881-490-162
Perfect score: 48
Sequence: 1 KKWLIQLKK 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	48	100.0	9	17	AAW04030	Antifungal peptide
2	48	100.0	9	18	AAW44559	Anti-fungal peptid
3	48	100.0	9	18	AAW44561	Anti-fungal peptid
4	48	100.0	9	18	AAW44563	Anti-fungal peptid
5	48	100.0	9	18	AAW44572	Anti-fungal peptid
6	48	100.0	9	18	AAW43753	Bactericidal/perme
7	48	100.0	9	18	AAW43749	Bactericidal/perme
8	48	100.0	9	20	AAY00536	Antifungal peptide
9	48	100.0	9	20	AAY00538	Antifungal peptide
10	48	100.0	9	20	AAY00540	Antifungal peptide
11	48	100.0	9	20	AAY00549	Antifungal peptide
12	48	100.0	9	21	AAY88384	Antifungal peptide
13	48	100.0	9	22	AAB65460	Anti-fungal peptid
14	48	100.0	9	22	AAB65462	Anti-fungal peptid
15	48	100.0	9	22	AAB65464	Anti-fungal peptid
16	48	100.0	9	22	AAB65473	Anti-fungal peptid
17	48	100.0	10	17	AAW04039	Antifungal peptide
18	48	100.0	10	17	AAW04040	Antifungal peptide
19	48	100.0	10	18	AAW44578	Anti-fungal peptid
20	48	100.0	10	18	AAW44579	Anti-fungal peptid
21	48	100.0	10	18	AAW43761	Bactericidal/perme
22	48	100.0	10	18	AAW43760	Bactericidal/perme
23	48	100.0	10	20	AAY00555	Antifungal peptide
24	48	100.0	10	20	AAY00556	Antifungal peptide
25	48	100.0	10	22	AAB65479	Anti-fungal peptid
26	48	100.0	10	22	AAB65480	Anti-fungal peptid
27	48	100.0	11	17	AAW04041	Antifungal peptide
28	48	100.0	11	18	AAW44580	Anti-fungal peptid
29	48	100.0	11	18	AAW43762	Bactericidal/perme
30	48	100.0	11	20	AAY00557	Antifungal peptide
31	48	100.0	11	22	AAB65481	Anti-fungal peptid
32	43	89.6	8	17	AAW04027	Antifungal peptide
33	43	89.6	8	17	AAW04029	Antifungal peptide
34	43	89.6	8	18	AAW44556	Anti-fungal peptid
35	43	89.6	8	18	AAW44558	Anti-fungal peptid
36	43	89.6	8	18	AAW43748	Bactericidal/perme
37	43	89.6	8	20	AAY00535	Antifungal peptide
38	43	89.6	8	20	AAY00533	Antifungal peptide
39	43	89.6	8	22	AAB65457	Anti-fungal peptid
40	43	89.6	8	22	AAB65459	Anti-fungal peptid
41	43	89.6	9	17	AAW04031	Antifungal peptide
42	43	89.6	9	17	AAW04036	Antifungal peptide
43	43	89.6	9	17	AAW04037	Antifungal peptide
44	43	89.6	9	18	AAW44564	Anti-fungal peptid
45	43	89.6	9	18	AAW44565	Anti-fungal peptid

ALIGNMENTS

RESULT 1
AAW04030
ID AAW04030 standard; peptide; 9 AA.
XX
AC AAW04030;
XX
DT 04-NOV-1996 (first entry)
XX
DE Antifungal peptide XMP.327/XMP.331/XMP.333/XMP.342.
XX
KW Antifungal peptide; inhibitor; Domain III; polymorphonuclear leukocyte;
KW bactericidal/permeability-increasing protein; BPI; mammalian; PMN; fungi;
KW neutrophil; replication inhibitor; fungal infection; Aspergillus;
KW Cryptococcus; Candida; C.albicans; C.galabrat; C.krusei; C.lusitaniae;
KW C.parapsilosis; C.tropicalis; therapy.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "optionally acetylated, optionally D-form residue"
FT Misc-difference 2
FT /note= "optionally D-form residue"
FT Misc-difference 8
FT /note= "optionally D-form residue"
FT Modified-site 9
FT /note= "amidated, optionally D-form residue"
XX
PN WO9608509-A1.
XX
PD 21-MAR-1996.
XX

```
PF   20-JUL-1995;   95WO-US09262.
XX
PR   13-JAN-1995;   95US-0372105.
PR   15-SEP-1994;   94US-0306473.
XX
PA   (XOMA ) XOMA CORP.
XX
PI   Fadem MB,  Lim E,  Little RG;
XX
DR   WPI; 1996-179900/18.
XX
PT   Antifungal peptide(s) derived from Domain III of BPI protein  - used
PT   in vitro for killing or inhibiting replication of fungi, esp.
PT   Candida species
XX
PS   Claim 5; Page 154; 199pp; English.
XX
CC   AAW04000-W04160 represent antifungal peptides.  These sequences are
CC   based on (or derived from) Domain III of the
CC   bactericidal/permeability-increasing protein (BPI).  BPI is a protein
CC   that can be isolated from the granules of mammalian polymorphonuclear
CC   leukocytes (PMNs or neutrophils).  These antifungal peptides can be used
CC   for killing, or inhibiting replication of, fungi in vitro.  These
CC   sequences can also be used for treatment of a fungal infection involving
CC   fungi from the species Candida, Aspergillus and Cryptococcus.  The
CC   sequences are especially useful for treating C.albicans, C.galabrat,
CC   C.krusei, C.lusitaniae, C.parapsilosis and C.tropicalis infections.
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 48;  DB 17;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 9.3e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              !!!!!!!!!
Db          1 KKWLIQLKK 9


RESULT 2
AAW44559
ID   AAW44559 standard; peptide; 9 AA.
XX
AC   AAW44559;
XX
DT   27-APR-1998  (first entry)
XX
DE   Anti-fungal peptide #160 based on BPI protein (residues 142-169).
XX
KW   Anti-fungal peptide; bactericidal-permeability-increasing protein; BPI;
KW   polymorphonuclear leukocyte; fungicide.
XX
OS   Synthetic.
OS   Mammalia.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   9
FT                   /note= "C-terminal amide"
XX
PN   WO9704008-A1.
XX
PD   06-FEB-1997.
XX
PF   21-MAR-1996;   96WO-US03845.
XX
PR   20-JUL-1995;   95US-0504841.
XX
PA   (XOMA ) XOMA CORP.
XX
PI   Fadem MB,  Lim E,  Little RG;
XX
DR   WPI; 1997-132578/12.
XX
PT   Anti-fungal peptide(s) derived from or based on domain III of
PT   bactericidal-permeability-increasing protein - are used in vitro or
PT   in vivo as a fungicides
XX
PS   Claim 1; Page 191; 230pp; English.
XX
CC   This is a specifically claimed anti-fungal peptide which is based on
CC   domain III (amino acids 142-160) of bactericidal-permeability-increasing
CC   protein (BPI), isolated from the granules of mammalian polymorphonuclear
CC   leukocytes. It is used in compositions with diluents, carriers or
CC   adjuvants to treat fungal infections in patients. It may also be used in
CC   vitro to kill or inhibit the replication of fungi, such as in
CC   decontaminating fluids and sterilising medical and implant devices.
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 48;  DB 18;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 9.3e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              !!!!!!!!!
Db          1 KKWLIQLKK 9


RESULT 3
AAW44561
ID   AAW44561 standard; peptide; 9 AA.
XX
AC   AAW44561;
XX
DT   27-APR-1998  (first entry)
XX
DE   Anti-fungal peptide #162 based on BPI protein (residues 142-169).
XX
KW   Anti-fungal peptide; bactericidal-permeability-increasing protein; BPI;
KW   polymorphonuclear leukocyte; fungicide.
XX
OS   Synthetic.
OS   Mammalia.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N-terminal acetyl"
FT   Modified-site   9
FT                   /note= "C-terminal amide"
XX
PN   WO9704008-A1.
XX
PD   06-FEB-1997.
XX
PF   21-MAR-1996;   96WO-US03845.
XX
PR   20-JUL-1995;   95US-0504841.
XX
PA   (XOMA ) XOMA CORP.
XX
PI   Fadem MB,  Lim E,  Little RG;
XX
DR   WPI; 1997-132578/12.
XX
PT   Anti-fungal peptide(s) derived from or based on domain III of
PT   bactericidal-permeability-increasing protein - are used in vitro or
PT   in vivo as a fungicides
XX
PS   Claim 1; Page 192; 230pp; English.
XX
CC   This is a specifically claimed anti-fungal peptide which is based on
CC   domain III (amino acids 142-160) of bactericidal-permeability-increasing
CC   protein (BPI), isolated from the granules of mammalian polymorphonuclear
CC   leukocytes. It is used in compositions with diluents, carriers or
CC   adjuvants to treat fungal infections in patients. It may also be used in
CC   vitro to kill or inhibit the replication of fungi, such as in
```

```
CC   decontaminating fluids and sterilising medical and implant devices.
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 48;  DB 18;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 9.3e-05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


RESULT 4
AAW44563
ID   AAW44563 standard; peptide; 9 AA.
XX
AC   AAW44563;
XX
DT   27-APR-1998  (first entry)
XX
DE   Anti-fungal peptide #164 based on BPI protein (residues 142-169).
XX
KW   Anti-fungal peptide; bactericidal-permeability-increasing protein; BPI;
KW   polymorphonuclear leukocyte; fungicide.
XX
OS   Synthetic.
OS   Mammalia.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1
FT                   /note= "D-form residue"
FT   Modified-site   9
FT                   /note= "C-terminal amide"
XX
PN   WO9704008-A1.
XX
PD   06-FEB-1997.
XX
PF   21-MAR-1996;   96WO-US03845.
XX
PR   20-JUL-1995;   95US-0504841.
XX
PA   (XOMA ) XOMA CORP.
XX
PI   Fadem MB,  Lim E,  Little RG;
XX
DR   WPI; 1997-132578/12.
XX
PT   Anti-fungal peptide(s) derived from or based on domain III of
PT   bactericidal-permeability-increasing protein - are used in vitro or
PT   in vivo as a fungicides
XX
PS   Claim 1; Page 192; 230pp; English.
XX
CC   This is a specifically claimed anti-fungal peptide which is based on
CC   domain III (amino acids 142-160) of bactericidal-permeability-increasing
CC   protein (BPI), isolated from the granules of mammalian polymorphonuclear
CC   leukocytes. It is used in compositions with diluents, carriers or
CC   adjuvants to treat fungal infections in patients. It may also be used in
CC   vitro to kill or inhibit the replication of fungi, such as in
CC   decontaminating fluids and sterilising medical and implant devices.
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 48;  DB 18;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 9.3e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


RESULT 5
AAW44572
ID   AAW44572 standard; peptide; 9 AA.
XX
AC   AAW44572;
XX
DT   27-APR-1998  (first entry)
XX
DE   Anti-fungal peptide #173 based on BPI protein (residues 142-169).
XX
KW   Anti-fungal peptide; bactericidal-permeability-increasing protein; BPI;
KW   polymorphonuclear leukocyte; fungicide.
XX
OS   Synthetic.
OS   Mammalia.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1
FT                   /note= "D-form residue"
FT   Misc-difference 2
FT                   /note= "D-form residue"
FT   Misc-difference 8
FT                   /note= "D-form residue"
FT   Misc-difference 9
FT                   /note= "D-form residue, C-terminal amide"
XX
PN   WO9704008-A1.
XX
PD   06-FEB-1997.
XX
PF   21-MAR-1996;   96WO-US03845.
XX
PR   20-JUL-1995;   95US-0504841.
XX
PA   (XOMA ) XOMA CORP.
XX
PI   Fadem MB,  Lim E,  Little RG;
XX
DR   WPI; 1997-132578/12.
XX
PT   Anti-fungal peptide(s) derived from or based on domain III of
PT   bactericidal-permeability-increasing protein - are used in vitro or
PT   in vivo as a fungicides
XX
PS   Claim 1; Page 196; 230pp; English.
XX
CC   This is a specifically claimed anti-fungal peptide which is based on
CC   domain III (amino acids 142-160) of bactericidal-permeability-increasing
CC   protein (BPI), isolated from the granules of mammalian polymorphonuclear
CC   leukocytes. It is used in compositions with diluents, carriers or
CC   adjuvants to treat fungal infections in patients. It may also be used in
CC   vitro to kill or inhibit the replication of fungi, such as in
CC   decontaminating fluids and sterilising medical and implant devices.
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 48;  DB 18;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 9.3e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


RESULT 6
AAW43753
ID   AAW43753 standard; peptide; 9 AA.
XX
AC   AAW43753;
XX
DT   20-APR-1998  (first entry)
```

```
XX
DE   Bactericidal/permeability increasing peptide XMP.331.
XX
KW   Bactericidal/permeability increasing peptide; BPI; fusion protein;
KW   bacterial infection; fungal infection; endotoxin; heparin;
KW   angiogenesis; fungicidal; recombinant DNA; vector.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "Acetylated"
FT   Modified-site   9
FT                   /note= "Amidated"
XX
PN   WO9735009-A1.
XX
PD   25-SEP-1997.
XX
PF   18-MAR-1997;   97WO-US05287.
XX
PR   22-MAR-1996;   96US-0621803.
XX
PA   (XOMA ) XOMA CORP.
XX
PI   Better MD;
XX
DR   WPI; 1997-480215/44.
XX
PT   Recombinant production of bactericidal/permeability increasing
PT   protein - by expression as a fusion protein in microbial host cells,
PT   then cleaving the BPI peptide from the carrier
XX
PS   Claim 10; Page 127; 186pp; English.
XX
CC   A new recombinant DNA vector construct has been developed which encodes
CC   a fusion protein and is suitable for introduction into a bacterial host.
CC   The vector comprises: (a) DNA encoding at least one cationic
CC   bactericidal/permeability increasing peptide (BPI), (b) DNA encoding a
CC   carrier protein, and (c) DNA encoding an amino acid (aa) cleavage site
CC   located between (a) and (b). The present sequence represents a
CC   specifically claimed BPI peptide. The peptides have many uses including
CC   the treatment of bacterial and fungal infections. BPI peptides also
CC   bind to endotoxins and heparin, neutralising their effects. The
CC   peptides have further been shown to inhibit angiogenesis (partly due to
CC   heparin-binding activity). The fusion proteins have been found to be
CC   expressed in large amounts without significant proteolysis, and in some
CC   cases are actually secreted from the host cells. This allows the
CC   indirect production of anti-microbial BPI peptides in microbial hosts.
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 48;  DB 18;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 9.3e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


RESULT 7
AAW43749
ID   AAW43749 standard; peptide; 9 AA.
XX
AC   AAW43749;
XX
DT   20-APR-1998  (first entry)
XX
DE   Bactericidal/permeability increasing peptide XMP.327.
XX
KW   Bactericidal/permeability increasing peptide; BPI; fusion protein;
KW   bacterial infection; fungal infection; endotoxin; heparin;
KW   angiogenesis; fungicidal; recombinant DNA; vector.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   9
FT                   /note= "Amidated"
XX
PN   WO9735009-A1.
XX
PD   25-SEP-1997.
XX
PF   18-MAR-1997;   97WO-US05287.
XX
PR   22-MAR-1996;   96US-0621803.
XX
PA   (XOMA ) XOMA CORP.
XX
PI   Better MD;
XX
DR   WPI; 1997-480215/44.
XX
PT   Recombinant production of bactericidal/permeability increasing
PT   protein - by expression as a fusion protein in microbial host cells,
PT   then cleaving the BPI peptide from the carrier
XX
PS   Claim 10; Page 125; 186pp; English.
XX
CC   A new recombinant DNA vector construct has been developed which encodes
CC   a fusion protein and is suitable for introduction into a bacterial host.
CC   The vector comprises: (a) DNA encoding at least one cationic
CC   bactericidal/permeability increasing peptide (BPI), (b) DNA encoding a
CC   carrier protein, and (c) DNA encoding an amino acid (aa) cleavage site
CC   located between (a) and (b). The present sequence represents a
CC   specifically claimed BPI peptide. The peptides have many uses including
CC   the treatment of bacterial and fungal infections. BPI peptides also
CC   bind to endotoxins and heparin, neutralising their effects. The
CC   peptides have further been shown to inhibit angiogenesis (partly due to
CC   heparin-binding activity). The fusion proteins have been found to be
CC   expressed in large amounts without significant proteolysis, and in some
CC   cases are actually secreted from the host cells. This allows the
CC   indirect production of anti-microbial BPI peptides in microbial hosts.
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 48;  DB 18;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 9.3e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


RESULT 8
AAY00536
ID   AAY00536 standard; Peptide; 9 AA.
XX
AC   AAY00536;
XX
DT   07-MAY-1999  (first entry)
XX
DE   Antifungal peptide XMP.327.
XX
KW   Antifungal; BPI; bactericidal/permeability increasing protein;
KW   Candida infection.
XX
OS   Synthetic.
XX
PN   US5858974-A.
XX
```

PD 12-JAN-1999.
XX
PF 21-MAR-1996; 96US-0621259.
XX
PR 21-MAR-1996; 96US-0621259.
PR 20-JUL-1995; 95US-0504841.
XX
PA (XOMA) XOMA CORP.
XX
PI Fadem MB, Lim E, Little RG;
XX
DR WPI; 1999-119956/10.
XX
PT Antifungal peptides - comprising part of bactericidal or
PT permeability-increasing protein sequence or related sequence
XX
PS Disclosure; Columns 165-166; 132pp; English.
XX
CC New peptides are provided which are based on Domain III (amino acids
CC 142-169) of human bactericidal/permeability-increasing protein (BPI).
CC The peptides all have a C-terminal amide. More particularly, the Claims
CC relate to: (1) a peptide that has an amino acid sequence of human BPI
CC from position 148 to position 161 (KSKVGWLIQLFHKK) and variants of the
CC sequence having antifungal activity; and (2) an antifungal peptide
CC having 6-14 amino acids comprising (a) a core sequence selected from
CC LIQL, IQLF, WLIQL, LIQLF and WLIQLF and (b) one or more cationic amino
CC acids selected from K, R, H, Orn (ornithine) and Dab (diaminobutyric
CC acid) at the N and/or C terminus of the core sequence. The new peptides
CC are used for killing or inhibiting replication of fungi in vitro; and
CC for treating fungal infections in vivo, in particular infections of
CC Candida, Aspergillus or Cryptococcus spp., especially C. albicans, C.
CC krusei, C. lusitaniae, C. parapsilosis or C. tropicalis. The peptide
CC can be administered topically, intravenously, orally or as an aerosol,
CC optionally together with a non-peptide antifungal agent.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 48; DB 20; Length 9;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKWLIQLKK 9
|||||||||
Db 1 KKWLIQLKK 9

RESULT 9
AAY00538
ID AAY00538 standard; Peptide; 9 AA.
XX
AC AAY00538;
XX
DT 07-MAY-1999 (first entry)
XX
DE Antifungal peptide XMP.331.
XX
KW Antifungal; BPI; bactericidal/permeability increasing protein;
KW Candida infection.
XX
OS Synthetic.
XX
PN US5858974-A.
XX
PD 12-JAN-1999.
XX
PF 21-MAR-1996; 96US-0621259.
XX
PR 21-MAR-1996; 96US-0621259.
PR 20-JUL-1995; 95US-0504841.
XX
PA (XOMA) XOMA CORP.
XX
PI Fadem MB, Lim E, Little RG;
XX
DR WPI; 1999-119956/10.
XX
PT Antifungal peptides - comprising part of bactericidal or
PT permeability-increasing protein sequence or related sequence
XX
PS Disclosure; Columns 165-166; 132pp; English.
XX
CC New peptides are provided which are based on Domain III (amino acids
CC 142-169) of human bactericidal/permeability-increasing protein (BPI).
CC The peptides all have a C-terminal amide. More particularly, the Claims
CC relate to: (1) a peptide that has an amino acid sequence of human BPI
CC from position 148 to position 161 (KSKVGWLIQLFHKK) and variants of the
CC sequence having antifungal activity; and (2) an antifungal peptide
CC having 6-14 amino acids comprising (a) a core sequence selected from
CC LIQL, IQLF, WLIQL, LIQLF and WLIQLF and (b) one or more cationic amino
CC acids selected from K, R, H, Orn (ornithine) and Dab (diaminobutyric
CC acid) at the N and/or C terminus of the core sequence. The new peptides
CC are used for killing or inhibiting replication of fungi in vitro; and
CC for treating fungal infections in vivo, in particular infections of
CC Candida, Aspergillus or Cryptococcus spp., especially C. albicans, C.
CC krusei, C. lusitaniae, C. parapsilosis or C. tropicalis. The peptide
CC can be administered topically, intravenously, orally or as an aerosol,
CC optionally together with a non-peptide antifungal agent.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 48; DB 20; Length 9;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKWLIQLKK 9
|||||||||
Db 1 KKWLIQLKK 9

RESULT 10
AAY00540
ID AAY00540 standard; Peptide; 9 AA.
XX
AC AAY00540;
XX
DT 07-MAY-1999 (first entry)
XX
DE Antifungal peptide XMP.333.
XX
KW Antifungal; BPI; bactericidal/permeability increasing protein;
KW Candida infection.
XX
OS Synthetic.
XX
PN US5858974-A.
XX
PD 12-JAN-1999.
XX
PF 21-MAR-1996; 96US-0621259.
XX
PR 21-MAR-1996; 96US-0621259.
PR 20-JUL-1995; 95US-0504841.
XX
PA (XOMA) XOMA CORP.
XX
PI Fadem MB, Lim E, Little RG;
XX
DR WPI; 1999-119956/10.
XX
PT Antifungal peptides - comprising part of bactericidal or
PT permeability-increasing protein sequence or related sequence
XX
PS Disclosure; Columns 167-168; 132pp; English.
XX
CC New peptides are provided which are based on Domain III (amino acids
CC 142-169) of human bactericidal/permeability-increasing protein (BPI).

CC The peptides all have a C-terminal amide. More particularly, the Claims
CC relate to: (1) a peptide that has an amino acid sequence of human BPI
CC from position 148 to position 161 (KSKVGWLIQLFHKK) and variants of the
CC sequence having antifungal activity; and (2) an antifungal peptide
CC having 6-14 amino acids comprising (a) a core sequence selected from
CC LIQL, IQLF, WLIQL, LIQLF and WLIQLF and (b) one or more cationic amino
CC acids selected from K, R, H, Orn (ornithine) and Dab (diaminobutyric
CC acid) at the N and/or C terminus of the core sequence. The new peptides
CC are used for killing or inhibiting replication of fungi in vitro; and
CC for treating fungal infections in vivo, in particular infections of
CC Candida, Aspergillus or Cryptococcus spp., especially C. albicans, C.
CC krusei, C. lusitaniae, C. parapsilosis or C. tropicalis. The peptide
CC can be administered topically, intravenously, orally or as an aerosol,
CC optionally together with a non-peptide antifungal agent.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 48; DB 20; Length 9;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKWLIQLKK 9
|||||||||
Db 1 KKWLIQLKK 9

RESULT 11
AAY00549
ID AAY00549 standard; Peptide; 9 AA.
XX
AC AAY00549;
XX
DT 07-MAY-1999 (first entry)
XX
DE Antifungal peptide XMP.342.
XX
KW Antifungal; BPI; bactericidal/permeability increasing protein;
KW Candida infection.
XX
OS Synthetic.
XX
PN US5858974-A.
XX
PD 12-JAN-1999.
XX
PF 21-MAR-1996; 96US-0621259.
XX
PR 21-MAR-1996; 96US-0621259.
PR 20-JUL-1995; 95US-0504841.
XX
PA (XOMA) XOMA CORP.
XX
PI Fadem MB, Lim E, Little RG;
XX
DR WPI; 1999-119956/10.
XX
PT Antifungal peptides - comprising part of bactericidal or
PT permeability-increasing protein sequence or related sequence
XX
PS Disclosure; Columns 173-174; 132pp; English.
XX
CC New peptides are provided which are based on Domain III (amino acids
CC 142-169) of human bactericidal/permeability-increasing protein (BPI).
CC The peptides all have a C-terminal amide. More particularly, the Claims
CC relate to: (1) a peptide that has an amino acid sequence of human BPI
CC from position 148 to position 161 (KSKVGWLIQLFHKK) and variants of the
CC sequence having antifungal activity; and (2) an antifungal peptide
CC having 6-14 amino acids comprising (a) a core sequence selected from
CC LIQL, IQLF, WLIQL, LIQLF and WLIQLF and (b) one or more cationic amino
CC acids selected from K, R, H, Orn (ornithine) and Dab (diaminobutyric
CC acid) at the N and/or C terminus of the core sequence. The new peptides
CC are used for killing or inhibiting replication of fungi in vitro; and
CC for treating fungal infections in vivo, in particular infections of
CC Candida, Aspergillus or Cryptococcus spp., especially C. albicans, C.
CC krusei, C. lusitaniae, C. parapsilosis or C. tropicalis. The peptide
CC can be administered topically, intravenously, orally or as an aerosol,
CC optionally together with a non-peptide antifungal agent.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 48; DB 20; Length 9;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKWLIQLKK 9
|||||||||
Db 1 KKWLIQLKK 9

RESULT 12
AAY88384
ID AAY88384 standard; peptide; 9 AA.
XX
AC AAY88384;
XX
DT 24-JUL-2000 (first entry)
XX
DE Antifungal peptide XMP.342.
XX
KW Antifungal peptide; XMP.342; BPI; microorganism; defence; antibacterial;
KW bactericidal/permeability increasing protein; antimicrobial; XMP. 445;
KW fungal infection; Candida; Aspergillus; Fusarium; growth inhibitor;
KW decontaminate; sterilize; medical equipment; prosthetic joint.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 1
FT /note= "D-form residue"
FT Misc-difference 2
FT /note= "D-form residue"
FT Misc-difference 8
FT /note= "D-form residue"
FT Modified-site 9
FT /note= "D-form residue, C-terminal amide"
XX
PN WO200018798-A1.
XX
PD 06-APR-2000.
XX
PF 24-SEP-1999; 99WO-US22277.
XX
PR 25-SEP-1998; 98US-0101958.
PR 25-NOV-1998; 98US-0109896.
XX
PA (XOMA) XOMA TECHNOLOGY LTD.
XX
PI Little RG;
XX
DR WPI; 2000-303442/26.
XX
PT A new antifungal peptide, designated XMP.445, used to treat fungal
PT infections is based on Domain III of
PT bactericidal/permeability-increasing protein -
XX
PS Example 2; Page 28; 31pp; English.
XX
CC This sequence represents an antifungal peptide designated XMP.342, which
CC is based on amino acids 153-157 of bactericidal/permeability increasing
CC (BPI) protein (see AAY88382). BPI is a protein isolated from the
CC granules of mammalian polymorphonuclear leukocytes (PMNs or neutrophils),
CC which are essential in the defence against invading microorganisms. BPI
CC protein products have antibacterial and antimicrobial activity. The
CC invention relates to antibacterial peptides, specifically XMP.445 (see
CC AAY88381). The peptide is used to treat fungal infections, particularly
CC infections of the species Candida, Aspergillus or Fusarium. The peptide

```
CC   can also be used to treat a bacterial infection, or to kill or inhibit
CC   the growth of bacteria in vitro. The peptide may be useful to treat
CC   animals or plants in vivo, and as an agent to decontaminate fluids and
CC   surfaces and to sterilize surgical or other medical equipment and
CC   implantable devices including prosthetic joints.
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 48;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 9.3e-05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


RESULT 13
AAB65460
ID   AAB65460 standard; Peptide; 9 AA.
XX
AC   AAB65460;
XX
DT   27-MAR-2001  (first entry)
XX
DE   Anti-fungal peptide XMP.327.
XX
KW   Human; BPI; antifungal; polymorphonuclear leukocyte; neutrophil;
KW   bactericidal/permeability-increasing protein; bactericidal;
KW   fungal infection.
XX
OS   Homo sapiens.
XX
PN   US6156730-A.
XX
PD   05-DEC-2000.
XX
PF   08-JAN-1999;   99US-0227659.
XX
PR   21-MAR-1996;   96US-0621259.
PR   12-MAR-1993;   93US-0030644.
PR   15-JUL-1993;   93US-0093202.
PR   14-JAN-1994;   94US-0183222.
PR   11-MAR-1994;   94US-0209762.
PR   11-JUL-1994;   94US-0273540.
PR   15-SEP-1994;   94US-0306473.
PR   13-JAN-1995;   95US-0372105.
PR   20-JUL-1995;   95US-0504841.
XX
PA   (XOMA ) XOMA CORP.
XX
PI   Lim E,  Fadem MB,  Little RG;
XX
DR   WPI; 2001-090160/10.
XX
PT   Novel anti-fungal peptides derived from domain III of
PT   bactericidal/permeability-increasing protein useful for killing or
PT   inhibiting replication of fungi and for treating fungal infections -
XX
PS   Example 2; Columns 169-170; 134pp; English.
XX
CC   The present invention relates to antifungal peptides (see
CC   AAB65301-B65550) derived from Domain III (amino acids 142-169) of
CC   bactericidal/permeability-increasing protein (BPI). The present sequence
CC   is one such antifungal peptide. BPI is a protein isolated from the
CC   granules of mammalian polymorphonuclear leukocytes (PMNs or
CC   neutrophils). BPI has potent bactericidal activity against a broad range
CC   of gram-negative bacteria. The peptides of the present invention are
CC   useful for killing or inhibiting replication of fungi, and treating
CC   infections caused by fungus belonging to Candida, Aspergillus,
CC   Cryptococcus species such as C.albicans, C.glabrata, C.krusei,
CC   C.lusitaniae, C.parapsilosis and C.tropicalis.
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 48;  DB 22;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 9.3e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


RESULT 14
AAB65462
ID   AAB65462 standard; Peptide; 9 AA.
XX
AC   AAB65462;
XX
DT   27-MAR-2001  (first entry)
XX
DE   Anti-fungal peptide XMP.331.
XX
KW   Human; BPI; antifungal; polymorphonuclear leukocyte; neutrophil;
KW   bactericidal/permeability-increasing protein; bactericidal;
KW   fungal infection.
XX
OS   Homo sapiens.
XX
PN   US6156730-A.
XX
PD   05-DEC-2000.
XX
PF   08-JAN-1999;   99US-0227659.
XX
PR   21-MAR-1996;   96US-0621259.
PR   12-MAR-1993;   93US-0030644.
PR   15-JUL-1993;   93US-0093202.
PR   14-JAN-1994;   94US-0183222.
PR   11-MAR-1994;   94US-0209762.
PR   11-JUL-1994;   94US-0273540.
PR   15-SEP-1994;   94US-0306473.
PR   13-JAN-1995;   95US-0372105.
PR   20-JUL-1995;   95US-0504841.
XX
PA   (XOMA ) XOMA CORP.
XX
PI   Lim E,  Fadem MB,  Little RG;
XX
DR   WPI; 2001-090160/10.
XX
PT   Novel anti-fungal peptides derived from domain III of
PT   bactericidal/permeability-increasing protein useful for killing or
PT   inhibiting replication of fungi and for treating fungal infections -
XX
PS   Example 2; Columns 171-172; 134pp; English.
XX
CC   The present invention relates to antifungal peptides (see
CC   AAB65301-B65550) derived from Domain III (amino acids 142-169) of
CC   bactericidal/permeability-increasing protein (BPI). The present sequence
CC   is one such antifungal peptide. BPI is a protein isolated from the
CC   granules of mammalian polymorphonuclear leukocytes (PMNs or
CC   neutrophils). BPI has potent bactericidal activity against a broad range
CC   of gram-negative bacteria. The peptides of the present invention are
CC   useful for killing or inhibiting replication of fungi, and treating
CC   infections caused by fungus belonging to Candida, Aspergillus,
CC   Cryptococcus species such as C.albicans, C.glabrata, C.krusei,
CC   C.lusitaniae, C.parapsilosis and C.tropicalis.
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 48;  DB 22;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 9.3e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


RESULT 15
AAB65464
ID   AAB65464 standard; Peptide; 9 AA.
XX
AC   AAB65464;
XX
DT   27-MAR-2001  (first entry)
XX
DE   Anti-fungal peptide XMP.333.
XX
KW   Human; BPI; antifungal; polymorphonuclear leukocyte; neutrophil;
KW   bactericidal/permeability-increasing protein; bactericidal;
KW   fungal infection.
XX
OS   Homo sapiens.
XX
PN   US6156730-A.
XX
PD   05-DEC-2000.
XX
PF   08-JAN-1999;   99US-0227659.
XX
PR   21-MAR-1996;   96US-0621259.
PR   12-MAR-1993;   93US-0030644.
PR   15-JUL-1993;   93US-0093202.
PR   14-JAN-1994;   94US-0183222.
PR   11-MAR-1994;   94US-0209762.
PR   11-JUL-1994;   94US-0273540.
PR   15-SEP-1994;   94US-0306473.
PR   13-JAN-1995;   95US-0372105.
PR   20-JUL-1995;   95US-0504841.
XX
PA   (XOMA ) XOMA CORP.
XX
PI   Lim E, Fadem MB, Little RG;
XX
DR   WPI; 2001-090160/10.
XX
PT   Novel anti-fungal peptides derived from domain III of
PT   bactericidal/permeability-increasing protein useful for killing or
PT   inhibiting replication of fungi and for treating fungal infections
XX
PS   Example 2; Columns 173-174; 134pp; English.
XX
CC   The present invention relates to antifungal peptides (see
CC   AAB65301-B65550) derived from Domain III (amino acids 142-169) of
CC   bactericidal/permeability-increasing protein (BPI). The present sequence
CC   is one such antifungal peptide. BPI is a protein isolated from the
CC   granules of mammalian polymorphonuclear leukocytes (PMNs or
CC   neutrophils). BPI has potent bactericidal activity against a broad range
CC   of gram-negative bacteria. The peptides of the present invention are
CC   useful for killing or inhibiting replication of fungi, and treating
CC   infections caused by fungus belonging to Candida, Aspergillus,
CC   Cryptococcus species such as C.albicans, C.glabrata, C.krusei,
CC   C.lusitaniae, C.parapsilosis and C.tropicalis.
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 48;  DB 22;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 9.3e-05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


Search completed: October 14, 2003, 14:41:31

Job time : 38.7917 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 14:33:23 ; Search time 12.75 Seconds
(without alignments)
29.866 Million cell updates/sec

Title: US-09-881-490-162
Perfect score: 48
Sequence: 1 KKWLIQLKK 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	48	100.0	9	2	US-08-621-803-193	Sequence 193, App
2	48	100.0	9	2	US-08-621-803-197	Sequence 197, App
3	48	100.0	9	2	US-08-621-259A-160	Sequence 160, App
4	48	100.0	9	2	US-08-621-259A-162	Sequence 162, App
5	48	100.0	9	2	US-08-621-259A-164	Sequence 164, App
6	48	100.0	9	2	US-08-621-259A-173	Sequence 173, App
7	48	100.0	9	3	US-09-217-352-193	Sequence 193, App
8	48	100.0	9	3	US-09-217-352-197	Sequence 197, App
9	48	100.0	9	5	PCT-US95-09262-160	Sequence 160, App
10	48	100.0	9	5	PCT-US95-09262-162	Sequence 162, App
11	48	100.0	9	5	PCT-US95-09262-164	Sequence 164, App
12	48	100.0	9	5	PCT-US95-09262-173	Sequence 173, App
13	48	100.0	10	2	US-08-621-803-204	Sequence 204, App
14	48	100.0	10	2	US-08-621-803-205	Sequence 205, App
15	48	100.0	10	2	US-08-621-259A-179	Sequence 179, App
16	48	100.0	10	2	US-08-621-259A-180	Sequence 180, App
17	48	100.0	10	3	US-09-217-352-204	Sequence 204, App
18	48	100.0	10	3	US-09-217-352-205	Sequence 205, App
19	48	100.0	10	5	PCT-US95-09262-179	Sequence 179, App
20	48	100.0	10	5	PCT-US95-09262-180	Sequence 180, App
21	48	100.0	11	2	US-08-621-803-206	Sequence 206, App
22	48	100.0	11	2	US-08-621-259A-181	Sequence 181, App
23	48	100.0	11	3	US-09-217-352-206	Sequence 206, App
24	48	100.0	11	5	PCT-US95-09262-181	Sequence 181, App
25	43	89.6	8	2	US-08-621-803-192	Sequence 192, App
26	43	89.6	8	2	US-08-621-259A-157	Sequence 157, App
27	43	89.6	8	2	US-08-621-259A-159	Sequence 159, App
28	43	89.6	8	3	US-09-217-352-192	Sequence 192, App
29	43	89.6	8	5	PCT-US95-09262-157	Sequence 157, App
30	43	89.6	8	5	PCT-US95-09262-159	Sequence 159, App
31	43	89.6	9	2	US-08-621-803-198	Sequence 198, App
32	43	89.6	9	2	US-08-621-803-202	Sequence 202, App
33	43	89.6	9	2	US-08-621-803-203	Sequence 203, App
34	43	89.6	9	2	US-08-621-259A-165	Sequence 165, App
35	43	89.6	9	2	US-08-621-259A-166	Sequence 166, App
36	43	89.6	9	2	US-08-621-259A-175	Sequence 175, App
37	43	89.6	9	2	US-08-621-259A-176	Sequence 176, App
38	43	89.6	9	2	US-08-621-259A-206	Sequence 206, App
39	43	89.6	9	2	US-08-621-259A-241	Sequence 241, App
40	43	89.6	9	2	US-08-621-259A-242	Sequence 242, App
41	43	89.6	9	3	US-09-217-352-198	Sequence 198, App
42	43	89.6	9	3	US-09-217-352-202	Sequence 202, App
43	43	89.6	9	3	US-09-217-352-203	Sequence 203, App
44	43	89.6	9	5	PCT-US95-09262-165	Sequence 165, App
45	43	89.6	9	5	PCT-US95-09262-166	Sequence 166, App

ALIGNMENTS

```
RESULT 1
US-08-621-803-193
; Sequence 193, Application US/08621803
; Patent No. 5851802
; GENERAL INFORMATION:
;   APPLICANT: Better, Marc D.
;   TITLE OF INVENTION: Methods for Recombinant Microbial Production of
;   TITLE OF INVENTION: Fusion Proteins and BPI-Derived Peptides
;   NUMBER OF SEQUENCES: 265
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
;     STREET: 6300 Sears Tower, 233 South Wacker Drive
;     CITY: Chicago
;     STATE: Illinois
;     COUNTRY: United States of America
;     ZIP: 60606-6402
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/621,803
;     FILING DATE: 22-MAR-1996
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Borun, Michael F.
;     REGISTRATION NUMBER: 25,447
;     REFERENCE/DOCKET NUMBER: 27129/33199
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 312/474-6300
;     TELEFAX: 312/474-0448
;     TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 193:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 9 amino acids
;     TYPE: amino acid
;     TOPOLOGY: linear
;   MOLECULE TYPE: peptide
;   FEATURE:
;     NAME/KEY: misc_feature
;     OTHER INFORMATION: "XMP.327"
;   FEATURE:
;     NAME/KEY: Modified-site
;     LOCATION: C-Terminus
;     OTHER INFORMATION: /label= Amidation
;     OTHER INFORMATION: /note= "The C-Terminus is Amidated."
US-08-621-803-193

  Query Match             100.0%;  Score 48;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.5e+05;
```

```
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


RESULT 2
US-08-621-803-197
; Sequence 197, Application US/08621803
; Patent No. 5851802
;  GENERAL INFORMATION:
;    APPLICANT:  Better, Marc D.
;    TITLE OF INVENTION:  Methods for Recombinant Microbial Production of
;    TITLE OF INVENTION:  Fusion Proteins and BPI-Derived Peptides
;    NUMBER OF SEQUENCES:  265
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,803
;      FILING DATE:  22-MAR-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25,447
;      REFERENCE/DOCKET NUMBER:  27129/33199
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  197:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.331"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1
;      OTHER INFORMATION:  /label= Acetylated
;      OTHER INFORMATION:  /note= "Position 1 is acetylated."
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-08-621-803-197

  Query Match             100.0%;  Score 48;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


RESULT 3
US-08-621-259A-160
; Sequence 160, Application US/08621259A
; Patent No. 5858974
;  GENERAL INFORMATION:
;    APPLICANT:  Little II, Roger G
;    APPLICANT:  Lim, Edward
;    APPLICANT:  Fadem, Mitchell B.
;    TITLE OF INVENTION:  Anti-Fungal Peptides
;    NUMBER OF SEQUENCES:  252
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McAndrews, Held & Malloy, Ltd.
;      STREET:  500 West Madison Street
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60661
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,259A
;      FILING DATE:  21-MAR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/504,841
;      FILING DATE:  20-JUL-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McNicholas, Janet M.
;      REGISTRATION NUMBER:  32,918
;      REFERENCE/DOCKET NUMBER:  11021US02
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/707-8889
;      TELEFAX:  312/707-9155
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  160:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.327"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-08-621-259A-160

  Query Match             100.0%;  Score 48;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.5e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


RESULT 4
US-08-621-259A-162
; Sequence 162, Application US/08621259A
; Patent No. 5858974
;  GENERAL INFORMATION:
;    APPLICANT:  Little II, Roger G
;    APPLICANT:  Lim, Edward
;    APPLICANT:  Fadem, Mitchell B.
;    TITLE OF INVENTION:  Anti-Fungal Peptides
;    NUMBER OF SEQUENCES:  252
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McAndrews, Held & Malloy, Ltd.
;      STREET:  500 West Madison Street
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
```

```
;       ZIP:  60661
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/621,259A
;       FILING DATE:  21-MAR-1996
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/504,841
;       FILING DATE:  20-JUL-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  McNicholas, Janet M.
;       REGISTRATION NUMBER:  32,918
;       REFERENCE/DOCKET NUMBER:  11021US02
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  312/707-8889
;       TELEFAX:  312/707-9155
;       TELEX:
;   INFORMATION FOR SEQ ID NO:  162:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  9 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
;     FEATURE:
;       NAME/KEY:  misc_feature
;       OTHER INFORMATION:  "XMP.331"
;     FEATURE:
;       NAME/KEY:  Modified-site
;       LOCATION:  1
;       OTHER INFORMATION:  /label= Acetylated
;       OTHER INFORMATION:  /note= "Position 1 is acetylated."
;     FEATURE:
;       NAME/KEY:  Modified-site
;       LOCATION:  C-Terminus
;       OTHER INFORMATION:  /label= Amidation
;       OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-08-621-259A-162

  Query Match             100.0%;  Score 48;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.5e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


RESULT 5
US-08-621-259A-164
; Sequence 164, Application US/08621259A
; Patent No. 5858974
;  GENERAL INFORMATION:
;    APPLICANT:  Little II, Roger G
;    APPLICANT:  Lim, Edward
;    APPLICANT:  Fadem, Mitchell B.
;    TITLE OF INVENTION:  Anti-Fungal Peptides
;    NUMBER OF SEQUENCES:  252
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McAndrews, Held & Malloy, Ltd.
;      STREET:  500 West Madison Street
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60661
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,259A
;      FILING DATE:  21-MAR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/504,841
;      FILING DATE:  20-JUL-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McNicholas, Janet M.
;      REGISTRATION NUMBER:  32,918
;      REFERENCE/DOCKET NUMBER:  11021US02
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/707-8889
;      TELEFAX:  312/707-9155
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  164:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.333"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1
;      OTHER INFORMATION:  /label= D-Lys
;      OTHER INFORMATION:  /note= "Position 1 is D-lysine"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-08-621-259A-164

  Query Match             100.0%;  Score 48;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.5e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


RESULT 6
US-08-621-259A-173
; Sequence 173, Application US/08621259A
; Patent No. 5858974
;  GENERAL INFORMATION:
;    APPLICANT:  Little II, Roger G
;    APPLICANT:  Lim, Edward
;    APPLICANT:  Fadem, Mitchell B.
;    TITLE OF INVENTION:  Anti-Fungal Peptides
;    NUMBER OF SEQUENCES:  252
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McAndrews, Held & Malloy, Ltd.
;      STREET:  500 West Madison Street
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60661
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,259A
;      FILING DATE:  21-MAR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/504,841
;      FILING DATE:  20-JUL-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McNicholas, Janet M.
```

```
;       REGISTRATION NUMBER:  32,918
;       REFERENCE/DOCKET NUMBER:  11021US02
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  312/707-8889
;       TELEFAX:  312/707-9155
;       TELEX:
;   INFORMATION FOR SEQ ID NO:  173:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  9 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
;     FEATURE:
;       NAME/KEY:  misc_feature
;       OTHER INFORMATION:  "XMP.342"
;     FEATURE:
;       NAME/KEY:  Modified-site
;       LOCATION:  1, 2, 8 & 9
;       OTHER INFORMATION:  /label= D-Lys
;       OTHER INFORMATION:  /note= "Positions 1, 2, 8 & 9 are D-lysine"
;     FEATURE:
;       NAME/KEY:  Modified-site
;       LOCATION:  C-Terminus
;       OTHER INFORMATION:  /label= Amidation
;       OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-08-621-259A-173

  Query Match             100.0%;  Score 48;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.5e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              !!!!!!!!!
Db          1 KKWLIQLKK 9


RESULT 7
US-09-217-352-193
; Sequence 193, Application US/09217352
; Patent No. 6274344
;  GENERAL INFORMATION:
;    APPLICANT:  Better, Marc D.
;    TITLE OF INVENTION:  Methods for Recombinant Microbial Production of
;    TITLE OF INVENTION:  Fusion Proteins and BPI-Derived Peptides
;    NUMBER OF SEQUENCES:  265
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/217,352
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/621,803
;      FILING DATE:  22-MAR-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25,447
;      REFERENCE/DOCKET NUMBER:  27129/33199
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  193:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.327"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-09-217-352-193

  Query Match             100.0%;  Score 48;  DB 3;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.5e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              !!!!!!!!!
Db          1 KKWLIQLKK 9


RESULT 8
US-09-217-352-197
; Sequence 197, Application US/09217352
; Patent No. 6274344
;  GENERAL INFORMATION:
;    APPLICANT:  Better, Marc D.
;    TITLE OF INVENTION:  Methods for Recombinant Microbial Production of
;    TITLE OF INVENTION:  Fusion Proteins and BPI-Derived Peptides
;    NUMBER OF SEQUENCES:  265
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/217,352
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/621,803
;      FILING DATE:  22-MAR-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25,447
;      REFERENCE/DOCKET NUMBER:  27129/33199
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  197:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.331"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1
;      OTHER INFORMATION:  /label= Acetylated
;      OTHER INFORMATION:  /note= "Position 1 is acetylated."
;    FEATURE:
```

```
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-09-217-352-197

  Query Match             100.0%;  Score 48;  DB 3;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.5e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


RESULT 9
PCT-US95-09262-160
; Sequence 160, Application PC/TUS9509262
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  Anti-Fungal Peptides
;    NUMBER OF SEQUENCES:  206
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/09262
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/372,105
;      FILING DATE:  13-JAN-95
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/306,473
;      FILING DATE:  15-SEP-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/273,540
;      FILING DATE:  11-JUL-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/209,762
;      FILING DATE:  11-MAR-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/183,222
;      FILING DATE:  14-JAN-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/093,202
;      FILING DATE:  15-JUL-93
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/030,644
;      FILING DATE:  12-MAR-93
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25,447
;      REFERENCE/DOCKET NUMBER:  27129/10040
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  160:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.327"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated"
PCT-US95-09262-160

  Query Match             100.0%;  Score 48;  DB 5;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.5e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


RESULT 10
PCT-US95-09262-162
; Sequence 162, Application PC/TUS9509262
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  Anti-Fungal Peptides
;    NUMBER OF SEQUENCES:  206
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/09262
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/372,105
;      FILING DATE:  13-JAN-95
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/306,473
;      FILING DATE:  15-SEP-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/273,540
;      FILING DATE:  11-JUL-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/209,762
;      FILING DATE:  11-MAR-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/183,222
;      FILING DATE:  14-JAN-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/093,202
;      FILING DATE:  15-JUL-93
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/030,644
;      FILING DATE:  12-MAR-93
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25,447
;      REFERENCE/DOCKET NUMBER:  27129/10040
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  162:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
```

```
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.331"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1
;      OTHER INFORMATION:  /label= Acetylated
;      OTHER INFORMATION:  /note= "Position 1 is acetylated"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated"
PCT-US95-09262-162

  Query Match             100.0%;  Score 48;  DB 5;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.5e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


RESULT 11
PCT-US95-09262-164
; Sequence 164, Application PC/TUS9509262
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  Anti-Fungal Peptides
;    NUMBER OF SEQUENCES:  206
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/09262
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/372,105
;      FILING DATE:  13-JAN-95
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/306,473
;      FILING DATE:  15-SEP-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/273,540
;      FILING DATE:  11-JUL-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/209,762
;      FILING DATE:  11-MAR-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/183,222
;      FILING DATE:  14-JAN-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/093,202
;      FILING DATE:  15-JUL-93
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/030,644
;      FILING DATE:  12-MAR-93
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25,447
;      REFERENCE/DOCKET NUMBER:  27129/10040
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  164:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.333"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1
;      OTHER INFORMATION:  /label= D-Lys
;      OTHER INFORMATION:  /note= "Position 1 is D-lysine"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated"
PCT-US95-09262-164

  Query Match             100.0%;  Score 48;  DB 5;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.5e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


RESULT 12
PCT-US95-09262-173
; Sequence 173, Application PC/TUS9509262
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  Anti-Fungal Peptides
;    NUMBER OF SEQUENCES:  206
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/09262
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/372,105
;      FILING DATE:  13-JAN-95
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/306,473
;      FILING DATE:  15-SEP-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/273,540
;      FILING DATE:  11-JUL-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/209,762
;      FILING DATE:  11-MAR-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/183,222
;      FILING DATE:  14-JAN-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/093,202
```

```
;      FILING DATE:  15-JUL-93
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/030,644
;      FILING DATE:  12-MAR-93
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25,447
;      REFERENCE/DOCKET NUMBER:  27129/10040
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  173:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.342"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1, 2, 8 & 9
;      OTHER INFORMATION:  /label= D-Lys
;      OTHER INFORMATION:  /note= "Positions 1, 2, 8 & 9 are D-lysine"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated"
PCT-US95-09262-173

  Query Match             100.0%;  Score 48;  DB 5;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


RESULT 13
US-08-621-803-204
; Sequence 204, Application US/08621803
; Patent No. 5851802
;  GENERAL INFORMATION:
;    APPLICANT:  Better, Marc D.
;    TITLE OF INVENTION:  Methods for Recombinant Microbial Production of
;    TITLE OF INVENTION:  Fusion Proteins and BPI-Derived Peptides
;    NUMBER OF SEQUENCES:  265
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,803
;      FILING DATE:  22-MAR-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25,447
;      REFERENCE/DOCKET NUMBER:  27129/33199
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  204:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.348"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-08-621-803-204

  Query Match             100.0%;  Score 48;  DB 2;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 0.24;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          2 KKWLIQLKK 10


RESULT 14
US-08-621-803-205
; Sequence 205, Application US/08621803
; Patent No. 5851802
;  GENERAL INFORMATION:
;    APPLICANT:  Better, Marc D.
;    TITLE OF INVENTION:  Methods for Recombinant Microbial Production of
;    TITLE OF INVENTION:  Fusion Proteins and BPI-Derived Peptides
;    NUMBER OF SEQUENCES:  265
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,803
;      FILING DATE:  22-MAR-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25,447
;      REFERENCE/DOCKET NUMBER:  27129/33199
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  205:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.349"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-08-621-803-205
```

```
  Query Match            100.0%;  Score 48;  DB 2;  Length 10;
  Best Local Similarity  100.0%;  Pred. No. 0.24;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


RESULT 15
US-08-621-259A-179
; Sequence 179, Application US/08621259A
; Patent No. 5858974
;  GENERAL INFORMATION:
;    APPLICANT:  Little II, Roger G
;    APPLICANT:  Lim, Edward
;    APPLICANT:  Fadem, Mitchell B.
;    TITLE OF INVENTION:  Anti-Fungal Peptides
;    NUMBER OF SEQUENCES:  252
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McAndrews, Held & Malloy, Ltd.
;      STREET:  500 West Madison Street
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60661
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,259A
;      FILING DATE:  21-MAR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/504,841
;      FILING DATE:  20-JUL-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McNicholas, Janet M.
;      REGISTRATION NUMBER:  32,918
;      REFERENCE/DOCKET NUMBER:  11021US02
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/707-8889
;      TELEFAX:  312/707-9155
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  179:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.348"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-08-621-259A-179

  Query Match            100.0%;  Score 48;  DB 2;  Length 10;
  Best Local Similarity  100.0%;  Pred. No. 0.24;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          2 KKWLIQLKK 10


Search completed: October 14, 2003, 14:42:45
Job time : 12.75 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 14:34:14 ; Search time 21.9375 Seconds
(without alignments)
66.104 Million cell updates/sec

Title: US-09-881-490-162
Perfect score: 48
Sequence: 1 KKWLIQLKK 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 600653 seqs, 161128416 residues

Total number of hits satisfying chosen parameters: 600653

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	48	100.0	9	9	US-09-765-527-193	Sequence 193, App
2	48	100.0	9	9	US-09-765-527-197	Sequence 197, App
3	48	100.0	9	9	US-09-881-490-160	Sequence 160, App
4	48	100.0	9	9	US-09-881-490-162	Sequence 162, App
5	48	100.0	9	9	US-09-881-490-164	Sequence 164, App
6	48	100.0	9	9	US-09-881-490-173	Sequence 173, App
7	48	100.0	10	9	US-09-765-527-204	Sequence 204, App
8	48	100.0	10	9	US-09-765-527-205	Sequence 205, App
9	48	100.0	10	9	US-09-881-490-179	Sequence 179, App
10	48	100.0	10	9	US-09-881-490-180	Sequence 180, App
11	48	100.0	11	9	US-09-765-527-206	Sequence 206, App
12	48	100.0	11	9	US-09-881-490-181	Sequence 181, App
13	43	89.6	8	9	US-09-765-527-192	Sequence 192, App
14	43	89.6	8	9	US-09-881-490-157	Sequence 157, App
15	43	89.6	8	9	US-09-881-490-159	Sequence 159, App
16	43	89.6	9	9	US-09-765-527-198	Sequence 198, App
17	43	89.6	9	9	US-09-765-527-202	Sequence 202, App
18	43	89.6	9	9	US-09-765-527-203	Sequence 203, App
19	43	89.6	9	9	US-09-881-490-165	Sequence 165, App
20	43	89.6	9	9	US-09-881-490-166	Sequence 166, App
21	43	89.6	9	9	US-09-881-490-175	Sequence 175, App
22	43	89.6	9	9	US-09-881-490-176	Sequence 176, App
23	43	89.6	10	9	US-09-765-527-216	Sequence 216, App
24	40	83.3	10	9	US-09-765-527-207	Sequence 207, App
25	40	83.3	10	9	US-09-881-490-182	Sequence 182, App
26	38	79.2	7	9	US-09-765-527-188	Sequence 188, App
27	38	79.2	7	9	US-09-765-527-189	Sequence 189, App
28	38	79.2	7	9	US-09-765-527-190	Sequence 190, App
29	38	79.2	7	9	US-09-765-527-191	Sequence 191, App
30	38	79.2	7	9	US-09-881-490-155	Sequence 155, App
31	38	79.2	7	9	US-09-881-490-156	Sequence 156, App
32	38	79.2	7	9	US-09-881-490-158	Sequence 158, App
33	38	79.2	9	9	US-09-765-527-236	Sequence 236, App
34	38	79.2	9	9	US-09-881-490-177	Sequence 177, App
35	38	79.2	10	9	US-09-765-527-214	Sequence 214, App
36	38	79.2	11	9	US-09-765-527-208	Sequence 208, App
37	38	79.2	11	9	US-09-881-490-183	Sequence 183, App
38	38	79.2	11	9	US-09-881-490-189	Sequence 189, App
39	37	77.1	10	9	US-09-765-527-218	Sequence 218, App
40	36	75.0	9	9	US-09-765-527-199	Sequence 199, App
41	36	75.0	9	9	US-09-881-490-167	Sequence 167, App
42	36	75.0	436	10	US-09-925-300-1640	Sequence 1640, Ap
43	35	72.9	55	11	US-09-764-891-3233	Sequence 3233, Ap
44	35	72.9	2073	11	US-09-978-244A-10	Sequence 10, Appl
45	34	70.8	9	9	US-09-765-527-201	Sequence 201, App

ALIGNMENTS

```
RESULT 1
US-09-765-527-193
; Sequence 193, Application US/09765527
; Patent No. US20020006638A1
;   GENERAL INFORMATION:
;        APPLICANT: Better, Marc D.
;        TITLE OF INVENTION: Methods for Recombinant Microbial Production of
;                            Fusion Proteins and BPI-Derived Peptides
;        NUMBER OF SEQUENCES: 265
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
;             STREET: 6300 Sears Tower, 233 South Wacker Drive
;             CITY: Chicago
;             STATE: Illinois
;             COUNTRY: United States of America
;             ZIP: 60606-6402
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/765,527
;             FILING DATE: 18-Jan-2001
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/621,803
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Borun, Michael F.
;             REGISTRATION NUMBER: 25,447
;             REFERENCE/DOCKET NUMBER: 27129/33199
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 312/474-6300
;             TELEFAX: 312/474-0448
;             TELEX: 25-3856
;   INFORMATION FOR SEQ ID NO: 193:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 9 amino acids
```

```
;             TYPE: amino acid
;             TOPOLOGY: linear
;         MOLECULE TYPE: peptide
;         FEATURE:
;             NAME/KEY:  misc_feature
;             OTHER INFORMATION:  "XMP.327"
;         FEATURE:
;             NAME/KEY:  Modified-site
;             LOCATION:  C-Terminus
;             OTHER INFORMATION:  /label= Amidation
;             /note= "The C-Terminus is Amidated."
;         SEQUENCE DESCRIPTION: SEQ ID NO: 193:
US-09-765-527-193

  Query Match             100.0%;  Score 48;  DB 9;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 5.4e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


RESULT 2
US-09-765-527-197
; Sequence 197, Application US/09765527
; Patent No. US20020006638A1
;   GENERAL INFORMATION:
;         APPLICANT: Better, Marc D.
;         TITLE OF INVENTION: Methods for Recombinant Microbial Production of
;                             Fusion Proteins and BPI-Derived Peptides
;         NUMBER OF SEQUENCES: 265
;         CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
;             STREET: 6300 Sears Tower, 233 South Wacker Drive
;             CITY: Chicago
;             STATE: Illinois
;             COUNTRY: United States of America
;             ZIP: 60606-6402
;         COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;         CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/765,527
;             FILING DATE: 18-Jan-2001
;         PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/621,803
;             FILING DATE: <Unknown>
;         ATTORNEY/AGENT INFORMATION:
;             NAME: Borun, Michael F.
;             REGISTRATION NUMBER: 25,447
;             REFERENCE/DOCKET NUMBER: 27129/33199
;         TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 312/474-6300
;             TELEFAX: 312/474-0448
;             TELEX: 25-3856
;    INFORMATION FOR SEQ ID NO: 197:
;         SEQUENCE CHARACTERISTICS:
;             LENGTH: 9 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;         MOLECULE TYPE: peptide
;         FEATURE:
;             NAME/KEY:  misc_feature
;             OTHER INFORMATION:  "XMP.331"
;         FEATURE:
;             NAME/KEY:  Modified-site
;             LOCATION:  1
;             OTHER INFORMATION:  /label= Acetylated
;             /note= "Position 1 is acetylated."
;         FEATURE:
;             NAME/KEY:  Modified-site
;             LOCATION:  C-Terminus
;             OTHER INFORMATION:  /label= Amidation
;             /note= "The C-Terminus is Amidated."
;         SEQUENCE DESCRIPTION: SEQ ID NO: 197:
US-09-765-527-197

  Query Match             100.0%;  Score 48;  DB 9;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 5.4e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


RESULT 3
US-09-881-490-160
; Sequence 160, Application US/09881490
; Patent No. US20020077298A1
;   GENERAL INFORMATION:
;         APPLICANT: Little II, Roger G.
;                    Lim, Edward
;                    Fadem, Mitchell B.
;         TITLE OF INVENTION: Anti-Fungal Peptides
;         NUMBER OF SEQUENCES: 211
;         CORRESPONDENCE ADDRESS:
;             ADDRESSEE: McAndrews, Held & Malloy, Ltd.
;             STREET: 500 West Madison Street, 34th FloorDrive
;             CITY: Chicago
;             STATE: Illinois
;             COUNTRY: United States of America
;             ZIP: 60661
;         COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;         CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/881,490
;             FILING DATE: 14-Jun-2001
;         PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/119,858
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: 08/372,105
;             FILING DATE: 13-JAN-95
;             APPLICATION NUMBER: 08/306,473
;             FILING DATE: 15-SEP-94
;             APPLICATION NUMBER: 08/273,540
;             FILING DATE: 11-JUL-94
;             APPLICATION NUMBER: 08/209,762
;             FILING DATE: 11-MAR-94
;             APPLICATION NUMBER: 08/183,222
;             FILING DATE: 14-JAN-94
;             APPLICATION NUMBER: 08/093,202
;             FILING DATE: 15-JUL-93
;             APPLICATION NUMBER: 08/030,644
;             FILING DATE: 12-MAR-93
;         ATTORNEY/AGENT INFORMATION:
;             NAME: McNicholas, Janet M.
;             REGISTRATION NUMBER: 32,918
;             REFERENCE/DOCKET NUMBER: 100-238/11021US01
;         TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 312/707-8889
;             TELEFAX: 312/707-9155
;             TELEX: 650 388-1248
;    INFORMATION FOR SEQ ID NO: 160:
;         SEQUENCE CHARACTERISTICS:
;             LENGTH: 9 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;         MOLECULE TYPE: peptide
;         FEATURE:
```

```
;             NAME/KEY:  misc_feature
;             OTHER INFORMATION:  "XMP.327"
;        FEATURE:
;             NAME/KEY:  Modified-site
;             LOCATION:  C-Terminus
;             OTHER INFORMATION:  /label= Amidation
;             /note= "The C-Terminus is Amidated"
;        SEQUENCE DESCRIPTION: SEQ ID NO: 160:
US-09-881-490-160

  Query Match             100.0%;  Score 48;  DB 9;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 5.4e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


RESULT 4
US-09-881-490-162
; Sequence 162, Application US/09881490
; Patent No. US20020077298A1
;   GENERAL INFORMATION:
;        APPLICANT: Little II, Roger G.
;                   Lim, Edward
;                   Fadem, Mitchell B.
;        TITLE OF INVENTION: Anti-Fungal Peptides
;        NUMBER OF SEQUENCES: 211
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: McAndrews, Held & Malloy, Ltd.
;             STREET: 500 West Madison Street, 34th FloorDrive
;             CITY: Chicago
;             STATE: Illinois
;             COUNTRY: United States of America
;             ZIP: 60661
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/881,490
;             FILING DATE: 14-Jun-2001
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/119,858
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: 08/372,105
;             FILING DATE: 13-JAN-95
;             APPLICATION NUMBER: 08/306,473
;             FILING DATE: 15-SEP-94
;             APPLICATION NUMBER: 08/273,540
;             FILING DATE: 11-JUL-94
;             APPLICATION NUMBER: 08/209,762
;             FILING DATE: 11-MAR-94
;             APPLICATION NUMBER: 08/183,222
;             FILING DATE: 14-JAN-94
;             APPLICATION NUMBER: 08/093,202
;             FILING DATE: 15-JUL-93
;             APPLICATION NUMBER: 08/030,644
;             FILING DATE: 12-MAR-93
;        ATTORNEY/AGENT INFORMATION:
;             NAME: McNicholas, Janet M.
;             REGISTRATION NUMBER: 32,918
;             REFERENCE/DOCKET NUMBER: 100-238/11021US01
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 312/707-8889
;             TELEFAX: 312/707-9155
;             TELEX: 650 388-1248
;   INFORMATION FOR SEQ ID NO: 162:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 9 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        FEATURE:
;             NAME/KEY:  misc_feature
;             OTHER INFORMATION:  "XMP.331"
;        FEATURE:
;             NAME/KEY:  Modified-site
;             LOCATION:  1
;             OTHER INFORMATION:  /label= Acetylated
;             /note= "Position 1 is acetylated"
;        FEATURE:
;             NAME/KEY:  Modified-site
;             LOCATION:  C-Terminus
;             OTHER INFORMATION:  /label= Amidation
;             /note= "The C-Terminus is Amidated"
;        SEQUENCE DESCRIPTION: SEQ ID NO: 162:
US-09-881-490-162

  Query Match             100.0%;  Score 48;  DB 9;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 5.4e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


RESULT 5
US-09-881-490-164
; Sequence 164, Application US/09881490
; Patent No. US20020077298A1
;   GENERAL INFORMATION:
;        APPLICANT: Little II, Roger G.
;                   Lim, Edward
;                   Fadem, Mitchell B.
;        TITLE OF INVENTION: Anti-Fungal Peptides
;        NUMBER OF SEQUENCES: 211
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: McAndrews, Held & Malloy, Ltd.
;             STREET: 500 West Madison Street, 34th FloorDrive
;             CITY: Chicago
;             STATE: Illinois
;             COUNTRY: United States of America
;             ZIP: 60661
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/881,490
;             FILING DATE: 14-Jun-2001
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/119,858
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: 08/372,105
;             FILING DATE: 13-JAN-95
;             APPLICATION NUMBER: 08/306,473
;             FILING DATE: 15-SEP-94
;             APPLICATION NUMBER: 08/273,540
;             FILING DATE: 11-JUL-94
;             APPLICATION NUMBER: 08/209,762
;             FILING DATE: 11-MAR-94
;             APPLICATION NUMBER: 08/183,222
;             FILING DATE: 14-JAN-94
;             APPLICATION NUMBER: 08/093,202
;             FILING DATE: 15-JUL-93
;             APPLICATION NUMBER: 08/030,644
;             FILING DATE: 12-MAR-93
;        ATTORNEY/AGENT INFORMATION:
;             NAME: McNicholas, Janet M.
;             REGISTRATION NUMBER: 32,918
;             REFERENCE/DOCKET NUMBER: 100-238/11021US01
```

```
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 312/707-8889
;             TELEFAX: 312/707-9155
;             TELEX: 650 388-1248
;   INFORMATION FOR SEQ ID NO: 164:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 9 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        FEATURE:
;             NAME/KEY:  misc_feature
;             OTHER INFORMATION:  "XMP.333"
;        FEATURE:
;             NAME/KEY:  Modified-site
;             LOCATION:  1
;             OTHER INFORMATION:  /label= D-Lys
;             /note= "Position 1 is D-lysine"
;        FEATURE:
;             NAME/KEY:  Modified-site
;             LOCATION:  C-Terminus
;             OTHER INFORMATION:  /label= Amidation
;             /note= "The C-Terminus is Amidated"
;        SEQUENCE DESCRIPTION: SEQ ID NO: 164:
US-09-881-490-164

  Query Match             100.0%;  Score 48;  DB 9;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 5.4e-05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


RESULT 6
US-09-881-490-173
; Sequence 173, Application US/09881490
; Patent No. US20020077298A1
;   GENERAL INFORMATION:
;        APPLICANT: Little II, Roger G.
;                   Lim, Edward
;                   Fadem, Mitchell B.
;        TITLE OF INVENTION: Anti-Fungal Peptides
;        NUMBER OF SEQUENCES: 211
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: McAndrews, Held & Malloy, Ltd.
;             STREET: 500 West Madison Street, 34th FloorDrive
;             CITY: Chicago
;             STATE: Illinois
;             COUNTRY: United States of America
;             ZIP: 60661
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/881,490
;             FILING DATE: 14-Jun-2001
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/119,858
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: 08/372,105
;             FILING DATE: 13-JAN-95
;             APPLICATION NUMBER: 08/306,473
;             FILING DATE: 15-SEP-94
;             APPLICATION NUMBER: 08/273,540
;             FILING DATE: 11-JUL-94
;             APPLICATION NUMBER: 08/209,762
;             FILING DATE: 11-MAR-94
;             APPLICATION NUMBER: 08/183,222
;             FILING DATE: 14-JAN-94
;             APPLICATION NUMBER: 08/093,202
;             FILING DATE: 15-JUL-93
;             APPLICATION NUMBER: 08/030,644
;             FILING DATE: 12-MAR-93
;        ATTORNEY/AGENT INFORMATION:
;             NAME: McNicholas, Janet M.
;             REGISTRATION NUMBER: 32,918
;             REFERENCE/DOCKET NUMBER: 100-238/11021US01
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 312/707-8889
;             TELEFAX: 312/707-9155
;             TELEX: 650 388-1248
;   INFORMATION FOR SEQ ID NO: 173:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 9 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        FEATURE:
;             NAME/KEY:  misc_feature
;             OTHER INFORMATION:  "XMP.342"
;        FEATURE:
;             NAME/KEY:  Modified-site
;             LOCATION:  1, 2, 8 & 9
;             OTHER INFORMATION:  /label= D-Lys
;             /note= "Positions 1, 2, 8 & 9 are D-lysine"
;        FEATURE:
;             NAME/KEY:  Modified-site
;             LOCATION:  C-Terminus
;             OTHER INFORMATION:  /label= Amidation
;             /note= "The C-Terminus is Amidated"
;        SEQUENCE DESCRIPTION: SEQ ID NO: 173:
US-09-881-490-173

  Query Match             100.0%;  Score 48;  DB 9;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 5.4e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


RESULT 7
US-09-765-527-204
; Sequence 204, Application US/09765527
; Patent No. US20020006638A1
;   GENERAL INFORMATION:
;        APPLICANT: Better, Marc D.
;        TITLE OF INVENTION: Methods for Recombinant Microbial Production of
;                            Fusion Proteins and BPI-Derived Peptides
;        NUMBER OF SEQUENCES: 265
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
;             STREET: 6300 Sears Tower, 233 South Wacker Drive
;             CITY: Chicago
;             STATE: Illinois
;             COUNTRY: United States of America
;             ZIP: 60606-6402
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/765,527
;             FILING DATE: 18-Jan-2001
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/621,803
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Borun, Michael F.
;             REGISTRATION NUMBER: 25,447
```

```
;            REFERENCE/DOCKET NUMBER: 27129/33199
;         TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 312/474-6300
;            TELEFAX: 312/474-0448
;            TELEX: 25-3856
;    INFORMATION FOR SEQ ID NO: 204:
;         SEQUENCE CHARACTERISTICS:
;            LENGTH: 10 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;         MOLECULE TYPE: peptide
;         FEATURE:
;            NAME/KEY: misc_feature
;            OTHER INFORMATION: "XMP.348"
;         FEATURE:
;            NAME/KEY: Modified-site
;            LOCATION: C-Terminus
;            OTHER INFORMATION: /label= Amidation
;            /note= "The C-Terminus is Amidated."
;         SEQUENCE DESCRIPTION: SEQ ID NO: 204:
US-09-765-527-204

  Query Match             100.0%;  Score 48;  DB 9;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 0.13;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          2 KKWLIQLKK 10


RESULT 8
US-09-765-527-205
; Sequence 205, Application US/09765527
; Patent No. US20020006638A1
;   GENERAL INFORMATION:
;        APPLICANT: Better, Marc D.
;        TITLE OF INVENTION: Methods for Recombinant Microbial Production of
;                            Fusion Proteins and BPI-Derived Peptides
;        NUMBER OF SEQUENCES: 265
;        CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
;            STREET: 6300 Sears Tower, 233 South Wacker Drive
;            CITY: Chicago
;            STATE: Illinois
;            COUNTRY: United States of America
;            ZIP: 60606-6402
;        COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/765,527
;            FILING DATE: 18-Jan-2001
;        PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 08/621,803
;            FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;            NAME: Borun, Michael F.
;            REGISTRATION NUMBER: 25,447
;            REFERENCE/DOCKET NUMBER: 27129/33199
;        TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 312/474-6300
;            TELEFAX: 312/474-0448
;            TELEX: 25-3856
;    INFORMATION FOR SEQ ID NO: 205:
;         SEQUENCE CHARACTERISTICS:
;            LENGTH: 10 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;         MOLECULE TYPE: peptide
;         FEATURE:
;            NAME/KEY: misc_feature
;            OTHER INFORMATION: "XMP.349"
;         FEATURE:
;            NAME/KEY: Modified-site
;            LOCATION: C-Terminus
;            OTHER INFORMATION: /label= Amidation
;            /note= "The C-Terminus is Amidated."
;         SEQUENCE DESCRIPTION: SEQ ID NO: 205:
US-09-765-527-205

  Query Match             100.0%;  Score 48;  DB 9;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 0.13;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


RESULT 9
US-09-881-490-179
; Sequence 179, Application US/09881490
; Patent No. US20020077298A1
;   GENERAL INFORMATION:
;        APPLICANT: Little II, Roger G.
;                   Lim, Edward
;                   Fadem, Mitchell B.
;        TITLE OF INVENTION: Anti-Fungal Peptides
;        NUMBER OF SEQUENCES: 211
;        CORRESPONDENCE ADDRESS:
;            ADDRESSEE: McAndrews, Held & Malloy, Ltd.
;            STREET: 500 West Madison Street, 34th FloorDrive
;            CITY: Chicago
;            STATE: Illinois
;            COUNTRY: United States of America
;            ZIP: 60661
;        COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/881,490
;            FILING DATE: 14-Jun-2001
;        PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/119,858
;            FILING DATE: <Unknown>
;            APPLICATION NUMBER: 08/372,105
;            FILING DATE: 13-JAN-95
;            APPLICATION NUMBER: 08/306,473
;            FILING DATE: 15-SEP-94
;            APPLICATION NUMBER: 08/273,540
;            FILING DATE: 11-JUL-94
;            APPLICATION NUMBER: 08/209,762
;            FILING DATE: 11-MAR-94
;            APPLICATION NUMBER: 08/183,222
;            FILING DATE: 14-JAN-94
;            APPLICATION NUMBER: 08/093,202
;            FILING DATE: 15-JUL-93
;            APPLICATION NUMBER: 08/030,644
;            FILING DATE: 12-MAR-93
;        ATTORNEY/AGENT INFORMATION:
;            NAME: McNicholas, Janet M.
;            REGISTRATION NUMBER: 32,918
;            REFERENCE/DOCKET NUMBER: 100-238/11021US01
;        TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 312/707-8889
;            TELEFAX: 312/707-9155
;            TELEX: 650 388-1248
;    INFORMATION FOR SEQ ID NO: 179:
;         SEQUENCE CHARACTERISTICS:
;            LENGTH: 10 amino acids
;            TYPE: amino acid
```

```
;           TOPOLOGY: linear
;       MOLECULE TYPE: peptide
;       FEATURE:
;           NAME/KEY:  misc_feature
;           OTHER INFORMATION:  "XMP.348"
;       FEATURE:
;           NAME/KEY:  Modified-site
;           LOCATION:  C-Terminus
;           OTHER INFORMATION:  /label= Amidation
;           /note= "The C-Terminus is Amidated"
;       SEQUENCE DESCRIPTION: SEQ ID NO: 179:
US-09-881-490-179

  Query Match             100.0%;  Score 48;  DB 9;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 0.13;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          2 KKWLIQLKK 10


RESULT 10
US-09-881-490-180
; Sequence 180, Application US/09881490
; Patent No. US20020077298A1
;   GENERAL INFORMATION:
;       APPLICANT: Little II, Roger G.
;                  Lim, Edward
;                  Fadem, Mitchell B.
;       TITLE OF INVENTION: Anti-Fungal Peptides
;       NUMBER OF SEQUENCES: 211
;       CORRESPONDENCE ADDRESS:
;           ADDRESSEE: McAndrews, Held & Malloy, Ltd.
;           STREET: 500 West Madison Street, 34th FloorDrive
;           CITY: Chicago
;           STATE: Illinois
;           COUNTRY: United States of America
;           ZIP: 60661
;       COMPUTER READABLE FORM:
;           MEDIUM TYPE: Floppy disk
;           COMPUTER: IBM PC compatible
;           OPERATING SYSTEM: PC-DOS/MS-DOS
;           SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;           APPLICATION NUMBER: US/09/881,490
;           FILING DATE: 14-Jun-2001
;       PRIOR APPLICATION DATA:
;           APPLICATION NUMBER: 09/119,858
;           FILING DATE: <Unknown>
;           APPLICATION NUMBER: 08/372,105
;           FILING DATE: 13-JAN-95
;           APPLICATION NUMBER: 08/306,473
;           FILING DATE: 15-SEP-94
;           APPLICATION NUMBER: 08/273,540
;           FILING DATE: 11-JUL-94
;           APPLICATION NUMBER: 08/209,762
;           FILING DATE: 11-MAR-94
;           APPLICATION NUMBER: 08/183,222
;           FILING DATE: 14-JAN-94
;           APPLICATION NUMBER: 08/093,202
;           FILING DATE: 15-JUL-93
;           APPLICATION NUMBER: 08/030,644
;           FILING DATE: 12-MAR-93
;       ATTORNEY/AGENT INFORMATION:
;           NAME: McNicholas, Janet M.
;           REGISTRATION NUMBER: 32,918
;           REFERENCE/DOCKET NUMBER: 100-238/11021US01
;       TELECOMMUNICATION INFORMATION:
;           TELEPHONE: 312/707-8889
;           TELEFAX: 312/707-9155
;           TELEX: 650 388-1248
;   INFORMATION FOR SEQ ID NO: 180:
;       SEQUENCE CHARACTERISTICS:
;           LENGTH: 10 amino acids
;           TYPE: amino acid
;           TOPOLOGY: linear
;       MOLECULE TYPE: peptide
;       FEATURE:
;           NAME/KEY:  misc_feature
;           OTHER INFORMATION:  "XMP.349"
;       FEATURE:
;           NAME/KEY:  Modified-site
;           LOCATION:  C-Terminus
;           OTHER INFORMATION:  /label= Amidation
;           /note= "The C-Terminus is Amidated"
;       SEQUENCE DESCRIPTION: SEQ ID NO: 180:
US-09-881-490-180

  Query Match             100.0%;  Score 48;  DB 9;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 0.13;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


RESULT 11
US-09-765-527-206
; Sequence 206, Application US/09765527
; Patent No. US20020006638A1
;   GENERAL INFORMATION:
;       APPLICANT: Better, Marc D.
;       TITLE OF INVENTION: Methods for Recombinant Microbial Production of
;                           Fusion Proteins and BPI-Derived Peptides
;       NUMBER OF SEQUENCES: 265
;       CORRESPONDENCE ADDRESS:
;           ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
;           STREET: 6300 Sears Tower, 233 South Wacker Drive
;           CITY: Chicago
;           STATE: Illinois
;           COUNTRY: United States of America
;           ZIP: 60606-6402
;       COMPUTER READABLE FORM:
;           MEDIUM TYPE: Floppy disk
;           COMPUTER: IBM PC compatible
;           OPERATING SYSTEM: PC-DOS/MS-DOS
;           SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;           APPLICATION NUMBER: US/09/765,527
;           FILING DATE: 18-Jan-2001
;       PRIOR APPLICATION DATA:
;           APPLICATION NUMBER: 08/621,803
;           FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;           NAME: Borun, Michael F.
;           REGISTRATION NUMBER: 25,447
;           REFERENCE/DOCKET NUMBER: 27129/33199
;       TELECOMMUNICATION INFORMATION:
;           TELEPHONE: 312/474-6300
;           TELEFAX: 312/474-0448
;           TELEX: 25-3856
;   INFORMATION FOR SEQ ID NO: 206:
;       SEQUENCE CHARACTERISTICS:
;           LENGTH: 11 amino acids
;           TYPE: amino acid
;           TOPOLOGY: linear
;       MOLECULE TYPE: peptide
;       FEATURE:
;           NAME/KEY:  misc_feature
;           OTHER INFORMATION:  "XMP.350"
;       FEATURE:
;           NAME/KEY:  Modified-site
;           LOCATION:  C-Terminus
;           OTHER INFORMATION:  /label= Amidation
```

```
;           /note= "The C-Terminus is Amidated."
;       SEQUENCE DESCRIPTION: SEQ ID NO: 206:
US-09-765-527-206

  Query Match             100.0%;  Score 48;  DB 9;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 0.14;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          2 KKWLIQLKK 10


RESULT 12
US-09-881-490-181
; Sequence 181, Application US/09881490
; Patent No. US20020077298A1
;   GENERAL INFORMATION:
;       APPLICANT: Little II, Roger G.
;                  Lim, Edward
;                  Fadem, Mitchell B.
;       TITLE OF INVENTION: Anti-Fungal Peptides
;       NUMBER OF SEQUENCES: 211
;       CORRESPONDENCE ADDRESS:
;           ADDRESSEE: McAndrews, Held & Malloy, Ltd.
;           STREET: 500 West Madison Street, 34th FloorDrive
;           CITY: Chicago
;           STATE: Illinois
;           COUNTRY: United States of America
;           ZIP: 60661
;       COMPUTER READABLE FORM:
;           MEDIUM TYPE: Floppy disk
;           COMPUTER: IBM PC compatible
;           OPERATING SYSTEM: PC-DOS/MS-DOS
;           SOFTWARE: Patentin Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;           APPLICATION NUMBER: US/09/881,490
;           FILING DATE: 14-Jun-2001
;       PRIOR APPLICATION DATA:
;           APPLICATION NUMBER: 09/119,858
;           FILING DATE: <Unknown>
;           APPLICATION NUMBER: 08/372,105
;           FILING DATE: 13-JAN-95
;           APPLICATION NUMBER: 08/306,473
;           FILING DATE: 15-SEP-94
;           APPLICATION NUMBER: 08/273,540
;           FILING DATE: 11-JUL-94
;           APPLICATION NUMBER: 08/209,762
;           FILING DATE: 11-MAR-94
;           APPLICATION NUMBER: 08/183,222
;           FILING DATE: 14-JAN-94
;           APPLICATION NUMBER: 08/093,202
;           FILING DATE: 15-JUL-93
;           APPLICATION NUMBER: 08/030,644
;           FILING DATE: 12-MAR-93
;       ATTORNEY/AGENT INFORMATION:
;           NAME: McNicholas, Janet M.
;           REGISTRATION NUMBER: 32,918
;           REFERENCE/DOCKET NUMBER: 100-238/11021US01
;       TELECOMMUNICATION INFORMATION:
;           TELEPHONE: 312/707-8889
;           TELEFAX: 312/707-9155
;           TELEX: 650 388-1248
;   INFORMATION FOR SEQ ID NO: 181:
;       SEQUENCE CHARACTERISTICS:
;           LENGTH: 11 amino acids
;           TYPE: amino acid
;           TOPOLOGY: linear
;       MOLECULE TYPE: peptide
;       FEATURE:
;           NAME/KEY:  misc_feature
;           OTHER INFORMATION:  "XMP.350"
;       FEATURE:
;           NAME/KEY:  Modified-site
;           LOCATION:  C-Terminus
;           OTHER INFORMATION:  /label= Amidation
;           /note= "The C-Terminus is Amidated"
;       SEQUENCE DESCRIPTION: SEQ ID NO: 181:
US-09-881-490-181

  Query Match             100.0%;  Score 48;  DB 9;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 0.14;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          2 KKWLIQLKK 10


RESULT 13
US-09-765-527-192
; Sequence 192, Application US/09765527
; Patent No. US20020006638A1
;   GENERAL INFORMATION:
;       APPLICANT: Better, Marc D.
;       TITLE OF INVENTION: Methods for Recombinant Microbial Production of
;                           Fusion Proteins and BPI-Derived Peptides
;       NUMBER OF SEQUENCES: 265
;       CORRESPONDENCE ADDRESS:
;           ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
;           STREET: 6300 Sears Tower, 233 South Wacker Drive
;           CITY: Chicago
;           STATE: Illinois
;           COUNTRY: United States of America
;           ZIP: 60606-6402
;       COMPUTER READABLE FORM:
;           MEDIUM TYPE: Floppy disk
;           COMPUTER: IBM PC compatible
;           OPERATING SYSTEM: PC-DOS/MS-DOS
;           SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;           APPLICATION NUMBER: US/09/765,527
;           FILING DATE: 18-Jan-2001
;       PRIOR APPLICATION DATA:
;           APPLICATION NUMBER: 08/621,803
;           FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;           NAME: Borun, Michael F.
;           REGISTRATION NUMBER: 25,447
;           REFERENCE/DOCKET NUMBER: 27129/33199
;       TELECOMMUNICATION INFORMATION:
;           TELEPHONE: 312/474-6300
;           TELEFAX: 312/474-0448
;           TELEX: 25-3856
;   INFORMATION FOR SEQ ID NO: 192:
;       SEQUENCE CHARACTERISTICS:
;           LENGTH: 8 amino acids
;           TYPE: amino acid
;           TOPOLOGY: linear
;       MOLECULE TYPE: peptide
;       FEATURE:
;           NAME/KEY:  misc_feature
;           OTHER INFORMATION:  "XMP.326"
;       FEATURE:
;           NAME/KEY:  Modified-site
;           LOCATION:  C-Terminus
;           OTHER INFORMATION:  /label= Amidation
;           /note= "The C-Terminus is Amidated."
;       SEQUENCE DESCRIPTION: SEQ ID NO: 192:
US-09-765-527-192

  Query Match             89.6%;  Score 43;  DB 9;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 5.4e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLK 8
```

```
              ||||||||
Db          1 KKWLIQLK 8


RESULT 14
US-09-881-490-157
; Sequence 157, Application US/09881490
; Patent No. US20020077298A1
;  GENERAL INFORMATION:
;       APPLICANT: Little II, Roger G.
;                  Lim, Edward
;                  Fadem, Mitchell B.
;       TITLE OF INVENTION: Anti-Fungal Peptides
;       NUMBER OF SEQUENCES: 211
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: McAndrews, Held & Malloy, Ltd.
;            STREET: 500 West Madison Street, 34th FloorDrive
;            CITY: Chicago
;            STATE: Illinois
;            COUNTRY: United States of America
;            ZIP: 60661
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/881,490
;            FILING DATE: 14-Jun-2001
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/119,858
;            FILING DATE: <Unknown>
;            APPLICATION NUMBER: 08/372,105
;            FILING DATE: 13-JAN-95
;            APPLICATION NUMBER: 08/306,473
;            FILING DATE: 15-SEP-94
;            APPLICATION NUMBER: 08/273,540
;            FILING DATE: 11-JUL-94
;            APPLICATION NUMBER: 08/209,762
;            FILING DATE: 11-MAR-94
;            APPLICATION NUMBER: 08/183,222
;            FILING DATE: 14-JAN-94
;            APPLICATION NUMBER: 08/093,202
;            FILING DATE: 15-JUL-93
;            APPLICATION NUMBER: 08/030,644
;            FILING DATE: 12-MAR-93
;       ATTORNEY/AGENT INFORMATION:
;            NAME: McNicholas, Janet M.
;            REGISTRATION NUMBER: 32,918
;            REFERENCE/DOCKET NUMBER: 100-238/11021US01
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 312/707-8889
;            TELEFAX: 312/707-9155
;            TELEX: 650 388-1248
;   INFORMATION FOR SEQ ID NO: 157:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 8 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: peptide
;       FEATURE:
;            NAME/KEY:  misc_feature
;            OTHER INFORMATION:  "XMP.324"
;       FEATURE:
;            NAME/KEY:  Modified-site
;            LOCATION:  C-Terminus
;            OTHER INFORMATION:  /label= Amidation
;            /note= "The C-Terminus is Amidated"
;       SEQUENCE DESCRIPTION: SEQ ID NO: 157:
US-09-881-490-157

  Query Match             89.6%;  Score 43;  DB 9;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 5.4e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 KWLIQLKK 9
              ||||||||
Db          1 KWLIQLKK 8


RESULT 15
US-09-881-490-159
; Sequence 159, Application US/09881490
; Patent No. US20020077298A1
;  GENERAL INFORMATION:
;       APPLICANT: Little II, Roger G.
;                  Lim, Edward
;                  Fadem, Mitchell B.
;       TITLE OF INVENTION: Anti-Fungal Peptides
;       NUMBER OF SEQUENCES: 211
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: McAndrews, Held & Malloy, Ltd.
;            STREET: 500 West Madison Street, 34th FloorDrive
;            CITY: Chicago
;            STATE: Illinois
;            COUNTRY: United States of America
;            ZIP: 60661
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/881,490
;            FILING DATE: 14-Jun-2001
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/119,858
;            FILING DATE: <Unknown>
;            APPLICATION NUMBER: 08/372,105
;            FILING DATE: 13-JAN-95
;            APPLICATION NUMBER: 08/306,473
;            FILING DATE: 15-SEP-94
;            APPLICATION NUMBER: 08/273,540
;            FILING DATE: 11-JUL-94
;            APPLICATION NUMBER: 08/209,762
;            FILING DATE: 11-MAR-94
;            APPLICATION NUMBER: 08/183,222
;            FILING DATE: 14-JAN-94
;            APPLICATION NUMBER: 08/093,202
;            FILING DATE: 15-JUL-93
;            APPLICATION NUMBER: 08/030,644
;            FILING DATE: 12-MAR-93
;       ATTORNEY/AGENT INFORMATION:
;            NAME: McNicholas, Janet M.
;            REGISTRATION NUMBER: 32,918
;            REFERENCE/DOCKET NUMBER: 100-238/11021US01
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 312/707-8889
;            TELEFAX: 312/707-9155
;            TELEX: 650 388-1248
;   INFORMATION FOR SEQ ID NO: 159:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 8 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: peptide
;       FEATURE:
;            NAME/KEY:  misc_feature
;            OTHER INFORMATION:  "XMP.326"
;       FEATURE:
;            NAME/KEY:  Modified-site
;            LOCATION:  C-Terminus
;            OTHER INFORMATION:  /label= Amidation
;            /note= "The C-Terminus is Amidated"
;       SEQUENCE DESCRIPTION: SEQ ID NO: 159:
US-09-881-490-159
```

Query Match 89.6%; Score 43; DB 9; Length 8;
Best Local Similarity 100.0%; Pred. No. 5.4e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 KKWLIQLK 8
              ||||||||
Db          1 KKWLIQLK 8
```

Search completed: October 14, 2003, 14:44:49
Job time : 21.9375 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 14:33:18 ; Search time 11.6125 Seconds
(without alignments)
73.271 Million cell updates/sec

Title: US-09-881-490-162
Perfect score: 48
Sequence: 1 KKWLIQLKK 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	37	77.1	189	2	AH1029	probable membrane
2	37	77.1	212	2	T31058	hypothetical prote
3	37	77.1	275	2	H81074	phytoene synthase-
4	36	75.0	70	2	G64500	hypothetical prote
5	36	75.0	89	2	F91003	Kil protein [impor
6	36	75.0	578	2	B71431	hypothetical prote
7	35	72.9	198	2	T51962	Rac-like GTP bindi
8	35	72.9	585	2	C70341	acetolactate synth
9	35	72.9	731	2	A99106	hypothetical prote
10	35	72.9	2655	2	D96595	probable acetyl-Co
11	34	70.8	59	2	AE2023	hypothetical prote
12	34	70.8	89	2	B85638	probable Kil prote
13	34	70.8	89	2	C95848	hypothetical prote
14	34	70.8	180	2	T37857	hypothetical prote
15	34	70.8	210	2	B40203	4-alpha-glucanotra
16	34	70.8	276	2	D81866	probable transfera
17	34	70.8	285	2	A69748	hypothetical prote
18	34	70.8	287	2	B90540	conserved hypothet
19	34	70.8	382	2	S55166	hypothetical prote
20	34	70.8	547	2	A64991	hypothetical ABC t
21	34	70.8	547	2	D91016	hypothetical prote
22	34	70.8	547	2	F85860	hypothetical prote
23	34	70.8	4588	2	T28667	dynein beta heavy
24	34	70.8	4639	1	A54794	dynein heavy chain
25	34	70.8	4725	1	A44357	dynein heavy chain
26	33	68.8	53	2	S72388	pd53 protein - Ent
27	33	68.8	95	2	I39575	nickel-cobalt-cadm
28	33	68.8	185	2	D82418	MutT/nudix family
29	33	68.8	185	2	A96989	transcription regu
30	33	68.8	201	2	T48862	rac-like protein A
31	33	68.8	209	2	T05164	hypothetical prote
32	33	68.8	220	2	AB0084	probable DedA-fami
33	33	68.8	233	2	D96763	hypothetical prote
34	33	68.8	252	2	F97072	ABC transporter, A
35	33	68.8	264	2	AE2787	exopolysaccharide
36	33	68.8	264	2	G97566	exopolysaccharide
37	33	68.8	335	2	A89834	teichoic acid bios
38	33	68.8	344	2	S46871	B13R protein - var
39	33	68.8	344	2	E72173	D2R protein - vari
40	33	68.8	344	2	T28611	hypothetical prote
41	33	68.8	367	2	AD3631	nitrous-oxide redu
42	33	68.8	396	2	T46901	hypothetical prote
43	33	68.8	585	2	A39286	parathyroid hormon
44	33	68.8	589	2	I59297	parathyroid hormon
45	33	68.8	591	2	S44203	parathyroid hormon

ALIGNMENTS

RESULT 1
AH1029
probable membrane protein STY4561 [imported] - Salmonella enterica subsp. enterica s
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AH1029
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churc
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Far
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica s
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AH1029
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-189 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD09337.1; PID:g16505337; GSPDB:GN00176
C;Genetics:
A;Gene: STY4561

Query Match 77.1%; Score 37; DB 2; Length 189;
Best Local Similarity 66.7%; Pred. No. 13;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 KKWLIQLKK 9
              |:|||| |:
Db        101 KRWLIQAKR 109
```

RESULT 2
T31058
hypothetical protein C05D9.5 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 17-Mar-2000
C;Accession: T31058
R;Fulton, L.; Gattung, S.
submitted to the EMBL Data Library, September 1999
A;Description: The sequence of C. elegans cosmid C05D9.
A;Reference number: Z20960
A;Accession: T31058
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-212 <FUL>
A;Cross-references: EMBL:U64605; PIDN:AAB04584.1
A;Experimental source: strain Bristol N2
C;Genetics:
A;Map position: X
A;Introns: 21/3; 82/1; 161/2
A;Note: C05D9.5
C;Superfamily: translation initiation factor eIF-4E

```
  Query Match             77.1%;  Score 37;  DB 2;  Length 212;
  Best Local Similarity   75.0%;  Pred. No. 14;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 KWLIQLKK 9
              ||:|:|||
Db        110 KWIIRLKK 117


RESULT 3
H81074
phytoene synthase-related protein NMB1521 [imported] - Neisseria meningitidis (strain MC
C;Species: Neisseria meningitidis
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C;Accession: H81074
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A;Reference number: A81000; MUID:20175755; PMID:10710307
A;Accession: H81074
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-275 <TET>
A;Cross-references: GB:AE002501; GB:AE002098; NID:g7226755; PIDN:AAF41877.1; PID:g722676
A;Experimental source: serogroup B, strain MC58
C;Genetics:
A;Gene: NMB1521

  Query Match             77.1%;  Score 37;  DB 2;  Length 275;
  Best Local Similarity   77.8%;  Pred. No. 19;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              | ||| |||
Db        263 KDWLIMLKK 271


RESULT 4
G64500
hypothetical protein MJ1608 - Methanococcus jannaschii
C;Species: Methanococcus jannaschii
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
C;Accession: G64500
R;Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,
; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A;Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C
A;Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
A;Reference number: A64300; MUID:96337999; PMID:8688087
A;Accession: G64500
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-70 <BUL>
A;Cross-references: GB:U67601; GB:L77117; NID:g2826439; PIDN:AAB99635.1; PID:g1500506; T
C;Genetics:
A;Map position: FOR1582918-1583130

  Query Match             75.0%;  Score 36;  DB 2;  Length 70;
  Best Local Similarity   75.0%;  Pred. No. 7.4;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLK 8
              |||:|:||
Db         63 KKWVIKLK 70


RESULT 5
F91003
Kil protein [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509952)
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 17-May-2002
C;Accession: F91003
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and g
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: F91003
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-89 <HAY>
A;Cross-references: GB:BA000007; PIDN:BAB36421.1; PID:g13362467; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs2998
C;Superfamily: phage lambda kil protein

  Query Match             75.0%;  Score 36;  DB 2;  Length 89;
  Best Local Similarity   55.6%;  Pred. No. 9.3;
  Matches    5;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||:: ||:
Db         74 KKWILMLKR 82


RESULT 6
B71431
hypothetical protein - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
A;Variety: columbia
C;Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 05-Dec-1998
C;Accession: B71431
R;Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; D
 P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenegger, T.; Pohl, T.M.; Terryn, N.; G
avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, E
Nature 391, 485-488, 1998
A;Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomer
erhoft, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.;
C.; Chalwatzis, N.
A;Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis t
A;Reference number: A71400; MUID:98121113; PMID:9461215
A;Accession: B71431
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-578 <BEV>
A;Cross-references: GB:Z97341; NID:g2244991; PID:e326992; PID:g2245000
C;Genetics:
A;Map position: 4COP9-4G3845

  Query Match             75.0%;  Score 36;  DB 2;  Length 578;
  Best Local Similarity   87.5%;  Pred. No. 59;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 KWLIQLKK 9
              ||| ||||
Db         30 KWLDQLKK 37


RESULT 7
T51962
Rac-like GTP binding protein [imported] - Picea mariana
C;Species: Picea mariana
C;Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 03-Nov-2000
C;Accession: T51962
R;Perry, D.J.; Bousquet, J.
Genetics 149, 1089-1098, 1998
A;Title: Sequence-tagged-site (STS) markers of arbitrary genes. Development, characte
A;Reference number: Z25268; MUID:98278823; PMID:9611216
A;Accession: T51962
A;Status: preliminary; translated from GB/EMBL/DDBJ
```

A;Molecule type: mRNA
A;Residues: 1-198 <PER>
A;Cross-references: EMBL:AF051223; PIDN:AAC32124.1
C;Genetics:
A;Note: Sb30
C;Superfamily: ras transforming protein; translation elongation factor Tu homology

Query Match 72.9%; Score 35; DB 2; Length 198;
Best Local Similarity 62.5%; Pred. No. 32;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKWLIQLK 8
||||::|:
Db 98 KKWLLELR 105

RESULT 8
C70341
acetolactate synthase (EC 4.1.3.18) large chain - Aquifex aeolicus
C;Species: Aquifex aeolicus
C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 15-Oct-1999
C;Accession: C70341
R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov
V.
Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A;Reference number: A70300; MUID:98196666; PMID:9537320
A;Accession: C70341
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-585 <AQF>
A;Cross-references: GB:AE000690; NID:g2983100; PIDN:AAC06706.1; PID:g2983101; GB:AE00065
A;Experimental source: strain VF5
C;Genetics:
A;Gene: ilvB
C;Superfamily: acetolactate synthase large chain; thiamin pyrophosphate-binding domain h
C;Keywords: branched-chain amino acid biosynthesis; carbon-carbon lyase; flavoprotein; m
F;440-488/Domain: thiamin pyrophosphate-binding domain homology <TPB>

Query Match 72.9%; Score 35; DB 2; Length 585;
Best Local Similarity 66.7%; Pred. No. 93;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KKWLIQLKK 9
|||| |::|
Db 351 KKWLEQIEK 359

RESULT 9
A99106
hypothetical protein orf731 [imported] - Guillardia theta nucleomorph
C;Species: nucleomorph Guillardia theta
A;Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001
C;Accession: A99106
R;Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Rei
Nature 410, 1091-1096, 2001
A;Title: The highly reduced genome of an enslaved algal nucleus.
A;Reference number: A99082; MUID:11323671; PMID:11323671
A;Accession: A99106
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-731 <DOU>
A;Cross-references: GB;AJ010592; NID:g12580677; PIDN:CAC26995.1; GSPDB:GN00151
C;Genetics:
A;Gene: orf731
A;Map position: 2
A;Genome: nucleomorph
C;Keywords: nucleomorph

Query Match 72.9%; Score 35; DB 2; Length 731;
Best Local Similarity 66.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KKWLIQLKK 9
|||: |||
Db 273 KKWIFLLKK 281

RESULT 10
D96595
probable acetyl-CoA synthetase, 45051-31547 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: D96595
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Al
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Ki
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marz
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tal
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: D96595
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-2655 <STO>
A;Cross-references: GB:AE005173; NID:g10645429; PIDN:AAG21546.1; GSPDB:GN00141
C;Genetics:
A;Gene: F7A10.14
A;Map position: 1

Query Match 72.9%; Score 35; DB 2; Length 2655;
Best Local Similarity 66.7%; Pred. No. 4.1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KKWLIQLKK 9
||| |||::
Db 2335 KKWPIQLRR 2343

RESULT 11
AE2023
hypothetical protein asl1739 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C;Accession: AE2023
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iri
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Taba
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacteriu
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AE2023
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-59 <KUR>
A;Cross-references: GB:BA000019; PIDN:BAB73438.1; PID:g17130829; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: asl1739

Query Match 70.8%; Score 34; DB 2; Length 59;
Best Local Similarity 55.6%; Pred. No. 15;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KKWLIQLKK 9
| ||::||:
Db 47 KAWLVKLKR 55

RESULT 12
B85638

probable Kil protein of bacteriphage BP-933W kilW [imported] - Escherichia coli (strain
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 17-May-2002
C;Accession: B85638
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: B85638
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-89 <STO>
A;Cross-references: GB:AE005174; NID:g12514292; PIDN:AAG55566.1; GSPDB:GN00145; UWGP:Z14
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: kilW
C;Superfamily: phage lambda kil protein

 Query Match 70.8%; Score 34; DB 2; Length 89;
 Best Local Similarity 62.5%; Pred. No. 22;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KKWLIQLK 8
||||:: ||
Db 74 KKWILMLK 81

RESULT 13
C85848
hypothetical protein Z3364 [imported] - Escherichia coli (strain O157:H7, substrain EDL9
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 17-May-2002
C;Accession: C85848
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: C85848
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-89 <STO>
A;Cross-references: GB:AE005174; NID:g12516417; PIDN:AAG57247.1; GSPDB:GN00145; UWGP:Z33
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: Z3364
C;Superfamily: phage lambda kil protein

 Query Match 70.8%; Score 34; DB 2; Length 89;
 Best Local Similarity 62.5%; Pred. No. 22;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KKWLIQLK 8
||||:: ||
Db 74 KKWILMLK 81

RESULT 14
T37857
hypothetical protein SPAC17G8.05 - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C;Accession: T37857
R;Devlin, R.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, February 1996
A;Reference number: Z21750
A;Accession: T37857
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-180 <DEV>
A;Cross-references: EMBL:Z69795; PIDN:CAA93688.1; GSPDB:GN00066; SPDB:SPAC17G8.05
A;Experimental source: strain 972h-; cosmid c17G8
C;Genetics:
A;Gene: SPDB:SPAC17G8.05
A;Map position: 1
A;Introns: 22/1

 Query Match 70.8%; Score 34; DB 2; Length 180;
 Best Local Similarity 62.5%; Pred. No. 44;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KKWLIQLK 8
||||::| |
Db 22 KKWVVQYK 29

RESULT 15
B40203
4-alpha-glucanotransferase (EC 2.4.1.25) / amylo-1,6-glucosidase (EC 3.2.1.33) - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 28-Aug-1992 #sequence_revision 28-Aug-1992 #text_change 16-Aug-2002
C;Accession: B40203
R;Yang, B.Z.; Ding, J.H.; Enghild, J.J.; Bao, Y.; Chen, Y.T.
J. Biol. Chem. 267, 9294-9299, 1992
A;Title: Molecular cloning and nucleotide sequence of cDNA encoding human muscle gly
A;Reference number: A40203; MUID:92250533; PMID:1374391
A;Accession: B40203
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-210 <YAN>
C;Superfamily: 4-alpha-glucanotransferase / amylo-1,6-glucosidase
C;Keywords: glycosidase; glycosyltransferase; hexosyltransferase; hydrolase; polysac

 Query Match 70.8%; Score 34; DB 2; Length 210;
 Best Local Similarity 62.5%; Pred. No. 52;
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 KWLIQLKK 9
|||::|:|
Db 203 KWLLELQK 210

Search completed: October 14, 2003, 14:38:03
Job time : 13.8125 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 14:33:18 ; Search time 6.375 Seconds
(without alignments)
66.391 Million cell updates/sec

Title: US-09-881-490-162
Perfect score: 48
Sequence: 1 KKWLIQLKK 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	37	77.1	212	1	IFE4_CAEEL	Q22888 caenorhabdi
2	36	75.0	70	1	YG08_METJA	Q59003 methanococc
3	34	70.8	180	1	YD65_SCHPO	Q10317 schizosacch
4	34	70.8	245	1	IFE3_HUMAN	O60573 h eukaryoti
5	34	70.8	245	1	IFE3_MOUSE	Q8bmb3 mus musculu
6	34	70.8	382	1	YJO7_YEAST	P47007 saccharomyc
7	34	70.8	430	1	SECY_STAAU	O08387 staphylococ
8	34	70.8	547	1	YOJI_ECOLI	P33941 escherichia
9	34	70.8	4639	1	DYHC_DROME	P37276 drosophila
10	34	70.8	4725	1	DYHC_DICDI	P34036 dictyosteli
11	33	68.8	95	1	NCCY_ALCXX	Q44581 alcaligenes
12	33	68.8	201	1	RAC2_ARATH	Q38903 arabidopsis
13	33	68.8	209	1	V282_ARATH	O65421 arabidopsis
14	33	68.8	218	1	GCH1_PASMU	P57865 pasteurella
15	33	68.8	267	1	EXOR_RHILV	Q52822 rhizobium l
16	33	68.8	344	1	SPI2_VARV	P33830 variola vir
17	33	68.8	396	1	PTD4_HUMAN	Q9ntk5 homo sapien
18	33	68.8	396	1	PTD4_MOUSE	Q9cz30 mus musculu
19	33	68.8	585	1	PTRR_DIDMA	P25107 didelphis m
20	33	68.8	585	1	PTRR_PIG	P50133 sus scrofa
21	33	68.8	591	1	PTRR_MOUSE	P41593 mus musculu
22	33	68.8	591	1	PTRR_RAT	P25961 rattus norv
23	33	68.8	593	1	PTRR_HUMAN	Q03431 homo sapien
24	33	68.8	639	1	NOSZ_RHIME	Q59746 rhizobium m
25	33	68.8	650	1	MI10_CAEEL	P34400 caenorhabdi
26	33	68.8	1742	1	MY5C_HUMAN	Q9nqx4 homo sapien
27	33	68.8	1877	1	PCK5_MOUSE	Q04592 mus musculu
28	33	68.8	2109	1	RRPL_VSVJH	P13615 vesicular s
29	32	66.7	290	1	Y084_MYCGE	P47330 mycoplasma
30	32	66.7	387	1	BODG_MOUSE	Q924y0 mus musculu
31	32	66.7	387	1	BODG_RAT	Q9qzu7 rattus norv
32	32	66.7	423	1	AATM_CHICK	P00508 gallus gall
33	32	66.7	454	1	Y095_MYCPN	P75538 mycoplasma
34	32	66.7	470	1	MRS2_YEAST	Q01926 saccharomyc
35	32	66.7	538	1	THIP_HAEIN	P44985 haemophilus
36	32	66.7	1067	1	CARB_CLOPE	Q8xhb3 clostridium
37	32	66.7	1328	1	POLX_TOBAC	P10978 nicotiana t
38	31.5	65.6	1953	1	BNI1_YEAST	P41832 saccharomyc
39	31	64.6	89	1	VKIL_LAMBD	P03758 bacteriopha
40	31	64.6	113	1	YB51_METJA	Q58551 methanococc
41	31	64.6	124	1	LBD9_ARATH	O82198 arabidopsis
42	31	64.6	152	1	FGF1_PIG	P20002 sus scrofa
43	31	64.6	155	1	FGF1_CHICK	P19596 gallus gall
44	31	64.6	155	1	FGF1_HUMAN	P05230 homo sapien
45	31	64.6	155	1	FGF1_MESAU	P34004 mesocricetu

ALIGNMENTS

RESULT 1
IFE4_CAEEL
ID IFE4_CAEEL STANDARD; PRT; 212 AA.
AC Q22888;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Eukaryotic translation initiation factor 4E (eIF-4E) (eIF4E) (mRNA
DE cap-binding protein) (eIF-4F 25 kDa subunit).
GN IFE-4 OR C05D9.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=20209441; PubMed=10744754;
RA Keiper B.D., Lamphear B.J., Deshpande A.M., Jankowska-Anyszka M.,
RA Aamodt E.J., Blumenthal T., Rhoads R.E.;
RT "Functional characterization of five eIF4E isoforms in Caenorhabditis
RT elegans.";
RL J. Biol. Chem. 275:10590-10596(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Fulton L., Gattung S.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DDBJ databases.
CC -!- FUNCTION: RECOGNIZES AND BINDS THE 7-METHYLGUANOSINE-CONTAINING
CC MRNA "CAP" DURING AN EARLY STEP IN THE INITIATION OF PROTEIN
CC SYNTHESIS AND FACILITATES RIBOSOME BINDING BY INDUCING THE
CC UNWINDING OF THE MRNAS SECONDARY STRUCTURES.
CC -!- SUBUNIT: EIF4F IS A TRIMER COMPOSED OF EIF4E, EIF4G AND EIF4A
CC (WHICH CAN CYCLE IN AND OUT OF THE COMPLEX).
CC -!- SIMILARITY: BELONGS TO THE EUKARYOTIC INITIATION FACTOR 4E FAMILY.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; AF214651; AAF62414.1; -.
DR EMBL; U64605; AAF98601.1; -.
DR PIR; T31058; T31058.
DR HSSP; P07260; 1AP8.
DR WormPep; C05D9.5; CE07931.
DR InterPro; IPR001040; TIF_eIF_4E.
DR Pfam; PF01652; IF4E; 1.
DR ProDom; PD003697; TIF_eIF_4E; 1.
DR PROSITE; PS00813; IF4E; 1.
KW Initiation factor; Protein biosynthesis; RNA-binding;
KW Multigene family.
SQ SEQUENCE 212 AA; 24584 MW; 87BA3DFE9FE8A486 CRC64;

```
  Query Match             77.1%;  Score 37;  DB 1;  Length 212;
  Best Local Similarity   75.0%;  Pred. No. 5.1;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 KWLIQLKK 9
              ||:|:|||
Db        110 KWIIRLKK 117


RESULT 2
YG08_METJA
ID   YG08_METJA     STANDARD;      PRT;    70 AA.
AC   Q59003;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical protein MJ1608.
GN   MJ1608.
OS   Methanococcus jannaschii.
OC   Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC   Methanocaldococcaceae; Methanocaldococcus.
OX   NCBI_TaxID=2190;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX   MEDLINE=96337999; PubMed=8688087;
RA   Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA   Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA   Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA   Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA   Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA   Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA   Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA   Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT   "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT   jannaschii.";
RL   Science 273:1058-1073(1996).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U67601; AAB99635.1; -.
DR   PIR; G64500; G64500.
DR   TIGR; MJ1608; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   70 AA;  8122 MW;  BDC6BE5A1A47AF57 CRC64;

  Query Match             75.0%;  Score 36;  DB 1;  Length 70;
  Best Local Similarity   75.0%;  Pred. No. 2.7;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLK 8
              |||:|:||
Db         63 KKWVIKLK 70


RESULT 3
YD65_SCHPO
ID   YD65_SCHPO     STANDARD;      PRT;    180 AA.
AC   Q10317;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hypothetical protein C17G8.05 in chromosome I.
GN   SPAC17G8.05.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RX   MEDLINE=21848401; PubMed=11859360;
RA   Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA   Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA   Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA   Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA   Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA   Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA   James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA   Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA   Oliver K., O'Neil S., Pearson D., Quail M.A., Rabbinowitsch E.,
RA   Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA   Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA   Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA   Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA   Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA   Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
RA   Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA   Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA   Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA   Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA   Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA   Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA   Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA   Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA   Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT   "The genome sequence of Schizosaccharomyces pombe.";
RL   Nature 415:871-880(2002).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z69795; CAA93688.1; -.
DR   PIR; T37857; T37857.
DR   GeneDB_SPombe; SPAC17G8.05; -.
KW   Hypothetical protein.
SQ   SEQUENCE   180 AA;  20931 MW;  A471D5243297E0B2 CRC64;

  Query Match             70.8%;  Score 34;  DB 1;  Length 180;
  Best Local Similarity   62.5%;  Pred. No. 16;
  Matches    5;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLK 8
              |||::| |
Db         22 KKWVVQYK 29


RESULT 4
IFE3_HUMAN
ID   IFE3_HUMAN     STANDARD;      PRT;    245 AA.
AC   O60573; O75349;
DT   15-SEP-2003 (Rel. 42, Created)
DT   15-SEP-2003 (Rel. 42, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Eukaryotic translation initiation factor 4E type 3 (eIF-4E type 3)
DE   (eIF4E type 3) (mRNA cap-binding protein type 3) (Eukaryotic
DE   translation initiation factor 4E-like 3) (Eukaryotic translation
DE   initiation factor 4E homologous protein) (mRNA cap-binding protein
DE   4EHP) (eIF4E-like protein 4E-LP).
GN   EIF4EL3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., CHARACTERIZATION, 3D-STRUCTURE MODELING, AND
RP   MUTAGENESIS.
RC   TISSUE=Follicle;
RX   MEDLINE=98250763; PubMed=9582349;
RA   Rom E., Kim H.C., Gingras A.-C., Marcotrigiano J., Favre D., Olsen H.,
RA   Burley S.K., Sonenberg N.;
RT   "Cloning and characterization of 4EHP, a novel mammalian eIF4E-related
RT   cap-binding protein.";
RL   J. Biol. Chem. 273:13104-13109(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Breast;
RA   Joshi B., Norris K.K., Jagus R.;
RT   "Isolation of human and mouse cDNAs encoding novel eIF4E-like
RT   proteins.";
RL   Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Blood;
RX   MEDLINE=98318631; PubMed=9653160;
RA   Mao M., Fu G., Wu J.-S., Zhang Q.-H., Zhou J., Kan L.-X., Huang Q.-H.,
RA   He K.-L., Gu B.-W., Han Z.-G., Shen Y., Gu J., Yu Y.-P., Xu S.-H.,
RA   Wang Y.-X., Chen S.-J., Chen Z.;
RT   "Identification of genes expressed in human CD34(+) hematopoietic
RT   stem/progenitor cells by expressed sequence tags and efficient full-
RT   length cDNA cloning.";
RL   Proc. Natl. Acad. Sci. U.S.A. 95:8175-8180(1998).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Muscle, Urinary bladder, and Uterus;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: Recognizes and binds the 7-methylguanosine-containing
CC       mRNA "CAP" during an early step in the initiation of protein
CC       synthesis and facilitates ribosome binding by inducing the
CC       unwinding of the mRNAS secondary structures.
CC   -!- SUBUNIT: EIF4F is a trimer composed of EIF4E, EIF4G and EIF4A
CC       (which can cycle in and out of the complex) (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE EUKARYOTIC INITIATION FACTOR 4E FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF047695; AAC18565.1; -.
DR   EMBL; AF068117; AAC19374.1; -.
DR   EMBL; AF038957; AAC39871.1; -.
DR   EMBL; BC005392; AAH05392.1; -.
DR   EMBL; BC005874; AAH05874.1; -.
DR   EMBL; BC021226; AAH21226.1; -.
DR   EMBL; BC021690; AAH21690.1; -.
DR   Genew; HGNC:3293; EIF4EL3.
DR   MIM; 605895; -.
DR   GO; GO:0003731; F:mRNA cap binding; TAS.
DR   GO; GO:0006445; P:regulation of translation; TAS.
DR   InterPro; IPR001040; TIF_eIF_4E.
DR   Pfam; PF01652; IF4E; 1.
DR   ProDom; PD003697; TIF_eIF_4E; 1.
DR   PROSITE; PS00813; IF4E; 1.
KW   Initiation factor; Protein biosynthesis; RNA-binding;
KW   Multigene family.
FT   MUTAGEN      63     63       W->A: UNABLE TO BIND CAPPED RNA.
FT   MUTAGEN      95     95       W->A: ABILITY TO BIND CAPPED MRNA REDUCED
FT                                TO 40% OF WILD-TYPE.
FT   MUTAGEN     124    126       WED->FAA: UNABLE TO BIND CAPPED RNA.
FT   MUTAGEN     124    124       W->A: ABILITY TO BIND CAPPED MRNA REDUCED
FT                                TO LESS THAN 10% OF WILD-TYPE.
FT   MUTAGEN     124    124       W->F: ABILITY TO BIND CAPPED MRNA REDUCED
FT                                TO 13% OF WILD-TYPE.
FT   MUTAGEN     125    125       E->A: ABILITY TO BIND CAPPED MRNA REDUCED
FT                                TO LESS THAN 10% OF WILD-TYPE.
FT   MUTAGEN     126    126       D->A: SLIGHT REDUCTION IN ABILITY TO BIND
FT                                CAPPED RNA.
FT   MUTAGEN     135    135       W->A: UNABLE TO BIND CAPPED RNA.
FT   MUTAGEN     148    148       W->A: UNABLE TO BIND CAPPED RNA.
FT   MUTAGEN     183    183       W->A: ABILITY TO BIND CAPPED MRNA REDUCED
FT                                TO LESS THAN 10% OF WILD-TYPE.
FT   MUTAGEN     183    183       W->F: UNABLE TO BIND CAPPED RNA.
FT   CONFLICT      1     27       MNNKFDALKDDDSGDHDQNEENSTQKD -> MMTVGTMIRM
FT                                KKTAHRKI (IN REF. 3).
SQ   SEQUENCE   245 AA;  28362 MW;  3D3075BFA48B3C12 CRC64;

  Query Match             70.8%;  Score 34;  DB 1;  Length 245;
  Best Local Similarity   62.5%;  Pred. No. 22;
  Matches    5;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 KWLIQLKK 9
              ||:|:|:|
Db        134 KWIIRLRK 141


RESULT 5
IFE3_MOUSE
ID   IFE3_MOUSE     STANDARD;      PRT;   245 AA.
AC   Q8BMB3; O88503;
DT   15-SEP-2003 (Rel. 42, Created)
DT   15-SEP-2003 (Rel. 42, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Eukaryotic translation initiation factor 4E type 3 (eIF-4E type 3)
DE   (eIF4E type 3) (mRNA cap-binding protein type 3) (Eukaryotic
DE   translation initiation factor 4E-like 3) (eIF4E-like protein 4E-LP).
GN   EIF4EL3.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NIH Swiss;
RA   Joshi B., Jagus R.;
RT   "Isolation of human and mouse cDNAs encoding novel eIF4E-like
RT   proteins.";
RL   Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Mesonephros;
RX   MEDLINE=22354683; PubMed=12466851;
RA   Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA   Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA   Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA   Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
```

```
RA   Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA   Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA   Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA   Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA   Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA   Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA   Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA   Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA   Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA   Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA   Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA   Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA   Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA   Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA   Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA   Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA   Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA   Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA   Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA   Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA   Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA   Birney E., Hayashizaki Y.;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
CC   -!- FUNCTION: Recognizes and binds the 7-methylguanosine-containing
CC       mRNA "CAP" during an early step in the initiation of protein
CC       synthesis and facilitates ribosome binding by inducing the
CC       unwinding of the mRNAS secondary structures (By similarity).
CC   -!- SUBUNIT: EIF4F is a trimer composed of EIF4E, EIF4G and EIF4A
CC       (which can cycle in and out of the complex) (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE EUKARYOTIC INITIATION FACTOR 4E FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF068116; AAC19373.1; -.
DR   EMBL; AK032965; BAC28102.1; -.
DR   MGD; MGI:1914440; Eif4el3.
DR   InterPro; IPR001040; TIF_eIF_4E.
DR   Pfam; PF01652; IF4E; 1.
DR   ProDom; PD003697; TIF_eIF_4E; 1.
DR   PROSITE; PS00813; IF4E; 1.
KW   Initiation factor; Protein biosynthesis; RNA-binding;
KW   Multigene family.
FT   CONFLICT    200    200       L -> F (IN REF. 1).
SQ   SEQUENCE   245 AA;  28263 MW;  AA722843721A01CC CRC64;

  Query Match             70.8%;  Score 34;  DB 1;  Length 245;
  Best Local Similarity   62.5%;  Pred. No. 22;
  Matches    5;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 KWLIQLKK 9
              |:|:|:|:|
Db        134 KWIIRLRK 141


RESULT 6
YJO7_YEAST
ID   YJO7_YEAST     STANDARD;      PRT;   382 AA.
AC   P47007;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   Hypothetical 44.9 kDa protein in INO1-IDS2 intergenic region.
GN   YJL147C OR J0639.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c / FY1679;
RX   MEDLINE=96408771; PubMed=8813765;
RA   Katsoulou C., Tzermia M., Tavernarakis N., Alexandraki D.;
RT   "Sequence analysis of a 40.7 kb segment from the left arm of yeast
RT   chromosome X reveals 14 known genes and 13 new open reading frames
RT   including homologues of genes clustered on the right arm of
RT   chromosome XI.";
RL   Yeast 12:787-797(1996).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z49422; CAA89442.1; -.
DR   EMBL; X87371; CAA60808.1; -.
DR   PIR; S55166; S55166.
DR   SGD; S0003683; YJL147C.
KW   Hypothetical protein.
SQ   SEQUENCE   382 AA;  44862 MW;  DDAAF588AF9A3234 CRC64;

  Query Match             70.8%;  Score 34;  DB 1;  Length 382;
  Best Local Similarity   66.7%;  Pred. No. 34;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||:|:|
Db        298 KKWLIDIVK 306


RESULT 7
SECY_STAAU
ID   SECY_STAAU     STANDARD;      PRT;   430 AA.
AC   O08387;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Preprotein translocase secY subunit.
GN   SECY.
OS   Staphylococcus aureus.
OC   Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX   NCBI_TaxID=1280;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NCTC 8325;
RA   Segarra R.A., Iandolo J.J.;
RL   Submitted (APR-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: INVOLVED IN PROTEIN EXPORT. INTERACTS WITH SECA AND SECE
CC       TO ALLOW THE TRANSLOCATION OF PROTEINS ACROSS THE PLASMA MEMBRANE,
CC       BY FORMING PART OF A CHANNEL.
CC   -!- SUBUNIT: ONE OF SEVEN SECRETORY PROTEINS (SECA-F & SECY) THAT
CC       COMPRISE THE PROKARYOTIC PROTEIN TRANSLOCATION APPARATUS.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO THE SECY/SEC61-ALPHA FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U96620; AAB54022.1; -.
DR   InterPro; IPR002208; SecY.
DR   Pfam; PF00344; secY; 1.
```

```
DR   PRINTS; PR00303; SECYTRNLCASE.
DR   TIGRFAMS; TIGR00967; 3a0501s007; 1.
DR   PROSITE; PS00755; SECY_1; 1.
DR   PROSITE; PS00756; SECY_2; 1.
KW   Protein transport; Translocation; Transmembrane.
SQ   SEQUENCE   430 AA;  47612 MW;  AF6923FE32434031 CRC64;

  Query Match             70.8%;  Score 34;  DB 1;  Length 430;
  Best Local Similarity   75.0%;  Pred. No. 38;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLK 8
              ||||| |:
Db        333 KKWLITLR 340


RESULT 8
YOJI_ECOLI
ID   YOJI_ECOLI     STANDARD;      PRT;   547 AA.
AC   P33941; P33942;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hypothetical ABC transporter ATP-binding protein yojI.
GN   YOJI OR B2211.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / BHB2600;
RA   Richterich P., Lakey N., Gryan G., Jaehn L., Mintz L., Robison K.,
RA   Church G.M.;
RL   Submitted (OCT-1993) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=97251358; PubMed=9097040;
RA   Itoh T., Aiba H., Baba T., Fujita K., Hayashi K., Inada T.,
RA   Isono K., Kasai H., Kimura S., Kitakawa M., Kitagawa M.,
RA   Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA   Nakade S., Nakamura Y., Nashimoto H., Nishio Y., Oshima T.,
RA   Saito N., Sampei G., Seki Y., Sivasundaram S., Tagami H.,
RA   Takeda J., Takemoto K., Wada C., Yamamoto Y., Horiuchi T.;
RT   "A 460-kb DNA sequence of the Escherichia coli K-12 genome
RT   corresponding to the 40.1-50.0 min region on the linkage map.";
RL   DNA Res. 3:379-392(1996).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC       (Probable).
CC   -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. SIMILAR TO
CC       P.SYRINGAE SYRD.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U00008; AAA16403.1; -.
DR   EMBL; AE000310; AAC75271.1; -.
DR   EMBL; D90850; BAA15994.1; -.
DR   PIR; A64991; A64991.
DR   EcoGene; EG12070; yojI.
DR   InterPro; IPR003593; AAA_ATPase.
DR   InterPro; IPR001140; ABC_TM_transpt.
DR   InterPro; IPR003439; ABC_transporter.
DR   InterPro; IPR005898; Cyc_pep_transpt.
DR   Pfam; PF00664; ABC_membrane; 1.
DR   Pfam; PF00005; ABC_tran; 1.
DR   ProDom; PD000006; ABC_transporter; 1.
DR   SMART; SM00382; AAA; 1.
DR   TIGRFAMS; TIGR01194; cyc_pep_trnsptr; 1.
DR   PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR   PROSITE; PS50893; ABC_TRANSPORTER_2; 1.
KW   Hypothetical protein; ATP-binding; Transport; Transmembrane;
KW   Inner membrane; Complete proteome.
FT   TRANSMEM     16     36       POTENTIAL.
FT   TRANSMEM     48     68       POTENTIAL.
FT   TRANSMEM    125    145       POTENTIAL.
FT   TRANSMEM    149    169       POTENTIAL.
FT   TRANSMEM    240    260       POTENTIAL.
FT   TRANSMEM    270    290       POTENTIAL.
FT   NP_BIND     356    363       ATP (POTENTIAL).
SQ   SEQUENCE   547 AA;  61552 MW;  11523935EDFD9EAB CRC64;

  Query Match             70.8%;  Score 34;  DB 1;  Length 547;
  Best Local Similarity   75.0%;  Pred. No. 48;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLK 8
              :||| |||
Db        426 EKWLAQLK 433


RESULT 9
DYHC_DROME
ID   DYHC_DROME     STANDARD;      PRT;  4639 AA.
AC   P37276;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Dynein heavy chain, cytosolic (DYHC).
GN   CDHC OR DHC64C.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94375524; PubMed=8089180;
RA   Li M., McGrail M., Serr M., Hays T.S.;
RT   "Drosophila cytoplasmic dynein, a microtubule motor that is
RT   asymmetrically localized in the oocyte.";
RL   J. Cell Biol. 126:1475-1494(1994).
RN   [2]
RP   SEQUENCE OF 1877-1998 FROM N.A.
RX   MEDLINE=94243034; PubMed=8186464;
RA   Rasmusson K., Serr M., Gepner J., Gibbons I., Hays T.S.;
RT   "A family of dynein genes in Drosophila melanogaster.";
RL   Mol. Biol. Cell 5:45-55(1994).
CC   -!- FUNCTION: Cytoplasmic dynein acts as a motor for the intracellular
CC       retrograde motility of vesicles and organelles along microtubules.
CC       Dynein has ATPase activity; the force-producing power stroke is
CC       thought to occur on release of ADP.
CC   -!- SUBUNIT: Consists of at least two heavy chains and a number of
CC       intermediate and light chains.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- DOMAIN: Dynein heavy chains probably consist of an N-terminal stem
CC       (which binds cargo and interacts with other dynein components),
CC       and the head or motor domain. The motor contains six tandemly-
CC       linked AAA domains in the head, which form a ring. A stalk-like
CC       structure (formed by two of the coiled coil domains) protrudes
```

```
CC       between AAA 4 and AAA 5 and terminates in a microtubule-binding
CC       site. A seventh domain may also contribute to this ring; it is not
CC       clear whether the N-terminus or the C-terminus forms this extra
CC       domain. There are four well-conserved and two non-conserved ATPase
CC       sites, one per AAA domain. Probably only one of these (within AAA
CC       1) actually hydrolyzes ATP, the others may serve a regulatory
CC       function.
CC   -!- SIMILARITY: Belongs to the dynein heavy chain family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L23195; AAA60323.1; -.
DR   EMBL; L25122; AAA28492.1; -.
DR   PIR; A54794; A54794.
DR   FlyBase; FBgn0010349; Dhc64C.
DR   GO; GO:0045169; C:fusome; NAS.
DR   GO; GO:0008567; F:dynein ATPase activity; IDA.
DR   GO; GO:0003777; F:microtubule motor activity; IDA.
DR   GO; GO:0007098; P:centrosome cycle; IMP.
DR   GO; GO:0045478; P:fusome organization and biogenesis; NAS.
DR   GO; GO:0007018; P:microtubule-based movement; IDA.
DR   GO; GO:0007294; P:oocyte determination; IMP.
DR   GO; GO:0007292; P:oogenesis; IMP.
DR   InterPro; IPR003593; AAA_ATPase.
DR   InterPro; IPR004273; Dynein_heavy.
DR   Pfam; PF03028; Dynein_heavy; 1.
DR   SMART; SM00382; AAA; 4.
KW   Motor protein; Dynein; Microtubules; ATP-binding; Repeat; Coiled coil.
FT   DOMAIN        1   1856       STEM (BY SIMILARITY).
FT   DOMAIN     1857   2084       AAA 1 (BY SIMILARITY).
FT   DOMAIN     2166   2437       AAA 2 (BY SIMILARITY).
FT   DOMAIN     2541   2790       AAA 3 (BY SIMILARITY).
FT   DOMAIN     2884   3153       AAA 4 (BY SIMILARITY).
FT   DOMAIN     3189   3478       STALK (BY SIMILARITY).
FT   DOMAIN     3539   3768       AAA 5 (BY SIMILARITY).
FT   DOMAIN     3989   4205       AAA 6 (BY SIMILARITY).
FT   DOMAIN      530    565       COILED COIL (POTENTIAL).
FT   DOMAIN      774    794       COILED COIL (POTENTIAL).
FT   DOMAIN     1264   1368       COILED COIL (POTENTIAL).
FT   DOMAIN     3189   3261       COILED COIL (POTENTIAL).
FT   DOMAIN     3382   3478       COILED COIL (POTENTIAL).
FT   DOMAIN     3723   3782       COILED COIL (POTENTIAL).
FT   NP_BIND    1895   1902       ATP (POTENTIAL).
FT   NP_BIND    2210   2217       ATP (POTENTIAL).
FT   NP_BIND    2580   2587       ATP (POTENTIAL).
FT   NP_BIND    2922   2929       ATP (POTENTIAL).
SQ   SEQUENCE   4639 AA;  530152 MW;  057A7D8800CCD07E CRC64;

  Query Match           70.8%;  Score 34;  DB 1;  Length 4639;
  Best Local Similarity   62.5%;  Pred. No. 3.8e+02;
  Matches    5;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 KWLIQLKK 9
              :|||:|||:|
Db       4088 QWLVQLEK 4095


RESULT 10
DYHC_DICDI
ID   DYHC_DICDI     STANDARD;      PRT;  4725 AA.
AC   P34036;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Dynein heavy chain, cytosolic (DYHC).
GN   DHCA.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AX3;
RX   MEDLINE=93107159; PubMed=1469051;
RA   Koonce M.P., Grissom P.M., McIntosh J.R.;
RT   "Dynein from Dictyostelium: primary structure comparisons between a
RT   cytoplasmic motor enzyme and flagellar dynein.";
RL   J. Cell Biol. 119:1597-1604(1992).
RN   [2]
RP   CHARACTERIZATION.
RX   MEDLINE=95170719; PubMed=7866389;
RA   Koonce M.P., Grissom P.M., Lyon M., Pope T., McIntosh J.R.;
RT   "Molecular characterization of a cytoplasmic dynein from
RT   Dictyostelium.";
RL   J. Eukaryot. Microbiol. 41:645-651(1994).
CC   -!- FUNCTION: Cytoplasmic dynein acts as a motor for the intracellular
CC       retrograde motility of vesicles and organelles along microtubules.
CC       Dynein has ATPase activity; the force-producing power stroke is
CC       thought to occur on release of ADP.
CC   -!- SUBUNIT: Consists of at least two heavy chains and a number of
CC       intermediate and light chains.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- DOMAIN: Dynein heavy chains probably consist of an N-terminal stem
CC       (which binds cargo and interacts with other dynein components),
CC       and the head or motor domain. The motor contains six tandemly-
CC       linked AAA domains in the head, which form a ring. A stalk-like
CC       structure (formed by two of the coiled coil domains) protrudes
CC       between AAA 4 and AAA 5 and terminates in a microtubule-binding
CC       site. A seventh domain may also contribute to this ring; it is not
CC       clear whether the N-terminus or the C-terminus forms this extra
CC       domain. There are four well-conserved and two non-conserved ATPase
CC       sites, one per AAA domain. Probably only one of these (within AAA
CC       1) actually hydrolyzes ATP, the others may serve a regulatory
CC       function.
CC   -!- SIMILARITY: Belongs to the dynein heavy chain family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z15124; CAA78827.1; -.
DR   PIR; A44357; A44357.
DR   DictyDb; DD01045; dhcA.
DR   InterPro; IPR003593; AAA_ATPase.
DR   InterPro; IPR004273; Dynein_heavy.
DR   Pfam; PF03028; Dynein_heavy; 1.
DR   SMART; SM00382; AAA; 3.
KW   Motor protein; Dynein; Microtubules; ATP-binding; Repeat; Coiled coil.
FT   DOMAIN        1   1930       STEM (BY SIMILARITY).
FT   DOMAIN     1931   2153       AAA 1 (BY SIMILARITY).
FT   DOMAIN     2233   2526       AAA 2 (BY SIMILARITY).
FT   DOMAIN     2630   2880       AAA 3 (BY SIMILARITY).
FT   DOMAIN     2973   3247       AAA 4 (BY SIMILARITY).
FT   DOMAIN     3266   3580       STALK (BY SIMILARITY).
FT   DOMAIN     3633   3862       AAA 5 (BY SIMILARITY).
FT   DOMAIN     4093   4307       AAA 6 (BY SIMILARITY).
FT   DOMAIN      864    897       COILED COIL (POTENTIAL).
FT   DOMAIN     1199   1219       COILED COIL (POTENTIAL).
FT   DOMAIN     1338   1367       COILED COIL (POTENTIAL).
FT   DOMAIN     1420   1436       COILED COIL (POTENTIAL).
FT   DOMAIN     1656   1684       COILED COIL (POTENTIAL).
FT   DOMAIN     2226   2248       COILED COIL (POTENTIAL).
FT   DOMAIN     2437   2457       COILED COIL (POTENTIAL).
FT   DOMAIN     3266   3344       COILED COIL (POTENTIAL).
FT   DOMAIN     3478   3580       COILED COIL (POTENTIAL).
FT   DOMAIN     3849   3876       COILED COIL (POTENTIAL).
FT   NP_BIND    1969   1976       ATP (POTENTIAL).
```

```
FT   NP_BIND    2271   2278       ATP (POTENTIAL).
FT   NP_BIND    2669   2676       ATP (POTENTIAL).
FT   NP_BIND    3011   3018       ATP (POTENTIAL).
SQ   SEQUENCE   4725 AA;  538807 MW;  9A0319915D579D3A CRC64;

  Query Match             70.8%;  Score 34;  DB 1;  Length 4725;
  Best Local Similarity   62.5%;  Pred. No. 3.8e+02;
  Matches    5;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 KWLIQLKK 9
              :||:||:|
Db       4190 QWLVQLEK 4197


RESULT 11
NCCY_ALCXX
ID   NCCY_ALCXX     STANDARD;      PRT;    95 AA.
AC   Q44581;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   Nickel-cobalt-cadmium resistance protein nccY.
GN   NCCY.
OS   Alcaligenes xylosoxydans xylosoxydans (Achromobacter xylosoxidans).
OG   Plasmid pTOM9.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC   Alcaligenaceae; Achromobacter.
OX   NCBI_TaxID=515;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=31A;
RX   MEDLINE=95050278; PubMed=7961470;
RA   Schmidt T., Schlegel H.G.;
RT   "Combined nickel-cobalt-cadmium resistance encoded by the ncc locus
RT   of Alcaligenes xylosoxidans 31A.";
RL   J. Bacteriol. 176:7045-7054(1994).
CC   -!- FUNCTION: COMPONENT OF THE NCC CATION-EFFLUX SYSTEM THAT CONFERS
CC       RESISTANCE TO NICKEL, COBALT AND CADMIUM. MAY BE INVOLVED IN THE
CC       REGULATION OF NCC.
CC   -!- SIMILARITY: TO A.EUTROPHUS CNRY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L31363; AAA65101.1; -.
DR   PIR; I39575; I39575.
KW   Plasmid; Nickel; Cobalt; Cadmium.
SQ   SEQUENCE   95 AA;  10422 MW;  4524AA1F3EB37987 CRC64;

  Query Match             68.8%;  Score 33;  DB 1;  Length 95;
  Best Local Similarity   62.5%;  Pred. No. 14;
  Matches    5;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 KWLIQLKK 9
              :||:| ||
Db          6 EWLVQAKK 13


RESULT 12
RAC2_ARATH
ID   RAC2_ARATH     STANDARD;      PRT;    201 AA.
AC   Q38903;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   RAC-like GTP binding protein ARAC2.
GN   ARAC2 OR AT5G45970 OR MCL19.1.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=98009984; PubMed=9349271;
RA   Winge P., Brembu T., Bones A.M.;
RT   "Cloning and characterization of rac-like cDNAs from Arabidopsis
RT   thaliana.";
RL   Plant Mol. Biol. 35:483-495(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Landsberg erecta;
RA   Winge P., Brembu T., Kristensen R., Bones A.M.;
RL   Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=98069011; PubMed=9405937;
RA   Kotani H., Nakamura Y., Sato S., Kaneko T., Asamizu E., Miyajima N.,
RA   Tabata S.;
RT   "Structural analysis of Arabidopsis thaliana chromosome 5. II.
RT   Sequence features of the regions of 1,044,062 bp covered by thirteen
RT   physically assigned P1 clones.";
RL   DNA Res. 4:291-300(1997).
CC   -!- TISSUE SPECIFICITY: EXPRESSED EXCLUSIVELY IN THE ROOT, HYPOCOTYL
CC       AND STEM.
CC   -!- SIMILARITY: BELONGS TO THE SMALL GTPASE SUPERFAMILY. RHO FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U43026; AAC49852.1; -.
DR   EMBL; AF115469; AAF40241.1; -.
DR   EMBL; AB006698; BAB08242.1; -.
DR   PIR; T48862; T48862.
DR   HSSP; P21181; 1AM4.
DR   InterPro; IPR003578; GTPase_Rho.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00174; RHO; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
KW   GTP-binding; Prenylation; Lipoprotein.
FT   NP_BIND      13     20       GTP (BY SIMILARITY).
FT   NP_BIND      60     64       GTP (BY SIMILARITY).
FT   NP_BIND     118    121       GTP (BY SIMILARITY).
FT   DOMAIN       35     43       EFFECTOR REGION (POTENTIAL).
FT   LIPID       198    198       GERANYL-GERANYL (POTENTIAL).
SQ   SEQUENCE   201 AA;  22393 MW;  B3E73F55BDBB8028 CRC64;

  Query Match             68.8%;  Score 33;  DB 1;  Length 201;
  Best Local Similarity   75.0%;  Pred. No. 28;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLK 8
              |||| :||
Db         98 KKWLPELK 105


RESULT 13
V282_ARATH
ID   V282_ARATH     STANDARD;      PRT;    209 AA.
AC   O65421;
DT   16-OCT-2001 (Rel. 40, Created)
```

```
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   VPS28 protein homolog 2.
GN   AT4G21560 OR F18E5.180 OR F17L22.20.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=20083488; PubMed=10617198;
RA   Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
RA   Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,
RA   Harris B., Ansorge W., Brandt P., Grivell L.A., Rieger M.,
RA   Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M.,
RA   Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidtheini T.,
RA   Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
RA   Vos P., Hoheisel J., Zimmermann W., Wedler H., Ridley P.,
RA   Langham S.-A., McCullagh B., Bilham L., Robben J.,
RA   Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandenbussche F.,
RA   Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
RA   Weitzenegger T., Bothe G., Ramsperger U., Hilbert H., Braun M.,
RA   Holzer E., Brandt A., Peters S., van Staveren M., Dirkse W.,
RA   Mooijman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
RA   Berneiser S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
RA   De Keyser A., Buysshaert C., Gielen J., Villarroel R., De Clercq R.,
RA   Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
RA   Clark L., Doggett J., Hall S., Kay M., Lennard N., McLay K., Mayes R.,
RA   Pettett A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,
RA   Borkova D., Bloecker H., Scharfe M., Grimm M., Loehnert T.-H.,
RA   Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
RA   Gabel C., Fuchs M., Fartmann B., Granderath K., Dauner D., Herzl A.,
RA   Neumann S., Argiriou A., Vitale D., Liguori R., Piravandi E.,
RA   Massenet O., Quigley F., Clabauld G., Muendlein A., Felber R.,
RA   Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
RA   Chefdor F., Cooke R., Berger C., Monfort A., Casacuberta E.,
RA   Gibbons T., Weber N., Vandenbol M., Bargues M., Terol J., Torres A.,
RA   Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
RA   Heijnen L., Schwarz S., Scholler P., Heber S., Francs P., Bielke C.,
RA   Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
RA   Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
RA   Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
RA   Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
RA   Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
RA   Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
RA   Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
RA   Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,
RA   Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
RA   Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshu C.,
RA   Antonoiu B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
RA   Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,
RA   Swaby I.K., O'Shaughnessy A., Rodriquez M., Hoffman J., Till S.,
RA   Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,
RA   Chen E., Marra M., Martienssen R., McCombie W.R.;
RT   "Sequence and analysis of chromosome 4 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 402:769-777(1999).
CC   -!- SIMILARITY: BELONGS TO THE VPS28 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AL022603; CAA18720.1; -.
DR   EMBL; AL035527; CAB36800.1; -.
DR   EMBL; AL161555; CAB81263.1; -.
DR   PIR; T05164; T05164.
DR   Pfam; PF03997; VPS28; 1.
KW   Hypothetical protein; Transport; Protein transport.
SQ   SEQUENCE   209 AA;  23494 MW;  8A8BB4B415D9786B CRC64;

  Query Match             68.8%;  Score 33;  DB 1;  Length 209;
  Best Local Similarity   55.6%;  Pred. No. 29;
  Matches    5;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |:||::| |
Db        166 KEWLLRLSK 174


RESULT 14
GCH1_PASMU
ID   GCH1_PASMU     STANDARD;      PRT;   218 AA.
AC   P57865;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   GTP cyclohydrolase I (EC 3.5.4.16) (GTP-CH-I).
GN   FOLE OR PM0693.
OS   Pasteurella multocida.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Pasteurella.
OX   NCBI_TaxID=747;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Pm70;
RX   MEDLINE=21145866; PubMed=11248100;
RA   May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT   "Complete genomic sequence of Pasteurella multocida Pm70.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC   -!- CATALYTIC ACTIVITY: GTP + 2 H(2)O = formate + 2-amino-4-hydroxy-6-
CC       (erythro-1,2,3-trihydroxypropyl)dihydropteridine triphosphate.
CC   -!- PATHWAY: Tetrahydrofolate biosynthesis; first step.
CC   -!- SUBUNIT: Homopolymer (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE GTP CYCLOHYDROLASE I FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE006106; AAK02777.1; -.
DR   HSSP; P27511; 1A8R.
DR   HAMAP; MF_00223; -; 1.
DR   InterPro; IPR001474; GTP_cyclohydroI.
DR   Pfam; PF01227; GTP_cyclohydroI; 1.
DR   ProDom; PD003330; GTP_cyclohydroI; 1.
DR   TIGRFAMs; TIGR00063; folE; 1.
DR   PROSITE; PS00859; GTP_CYCLOHYDROL_1_1; 1.
DR   PROSITE; PS00860; GTP_CYCLOHYDROL_1_2; 1.
KW   One-carbon metabolism; Hydrolase; Complete proteome.
FT   DISULFID    109    180       BY SIMILARITY.
SQ   SEQUENCE   218 AA;  24953 MW;  12F1FC8129E790C9 CRC64;

  Query Match             68.8%;  Score 33;  DB 1;  Length 218;
  Best Local Similarity   66.7%;  Pred. No. 30;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||:| | |
Db        127 KKWVIGLSK 135


RESULT 15
EXOR_RHILV
ID   EXOR_RHILV     STANDARD;      PRT;   267 AA.
AC   Q52822;
```

```
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Exopolysaccharide production negative regulator precursor.
GN   EXOR.
OS   Rhizobium leguminosarum (biovar viciae).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.
OX   NCBI_TaxID=387;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WSM710 / WR6-35;
RX   MEDLINE=97346048; PubMed=9202471;
RA   Reeve W.G., Dilworth M.J., Tiwari R.P., Glenn A.R.;
RT   "Regulation of exopolysaccharide production in Rhizobium leguminosarum
RT   biovar viciae WSM710 involves exoR.";
RL   Microbiology 143:1951-1958(1997).
CC   -!- FUNCTION: NEGATIVELY MODULATES EXOPOLYSACCHARIDE (EPS)
CC       BIOSYNTHESIS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; L39937; AAB63941.1; -.
DR   InterPro; IPR006597; Sel_like.
DR   SMART; SM00671; SEL1; 3.
KW   Exopolysaccharide synthesis; Signal.
FT   SIGNAL        1     23       POTENTIAL.
FT   CHAIN        24    267       EXOPOLYSACCHARIDE PRODUCTION NEGATIVE
FT                                REGULATOR.
SQ   SEQUENCE   267 AA;  28922 MW;  E45E2E0538E1F573 CRC64;

  Query Match             68.8%;  Score 33;  DB 1;  Length 267;
  Best Local Similarity   66.7%;  Pred. No. 37;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||| | :|
Db        185 KKWLNQARK 193


Search completed: October 14, 2003, 14:34:06
Job time : 8.54167 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 14:33:18 ; Search time 29.8125 Seconds
(without alignments)
77.903 Million cell updates/sec

Title: US-09-881-490-162
Perfect score: 48
Sequence: 1 KKWLIQLKK 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	40	83.3	369	2	Q93CS6	Q93cs6 shigella bo
2	38	79.2	229	5	Q8MP31	Q8mp31 dictyosteli
3	37	77.1	189	16	Q8Z1K3	Q8z1k3 salmonella
4	37	77.1	275	16	Q9JYL8	Q9jyl8 neisseria m
5	36	75.0	89	2	Q9EYA7	Q9eya7 escherichia
6	36	75.0	278	4	Q9UJ05	Q9uj05 homo sapien
7	36	75.0	284	4	Q96BV3	Q96bv3 homo sapien
8	36	75.0	459	10	Q8H3G5	Q8h3g5 oryza sativ
9	36	75.0	578	10	O23488	O23488 arabidopsis
10	36	75.0	652	5	Q8I427	Q8i427 plasmodium
11	36	75.0	1566	5	Q9N3E9	Q9n3e9 caenorhabdi
12	35	72.9	136	16	Q8EI76	Q8ei76 shewanella
13	35	72.9	198	10	O65062	O65062 picea maria
14	35	72.9	261	16	Q9CNC3	Q9cnc3 pasteurella
15	35	72.9	294	8	Q8HIS8	Q8his8 monosiga br
16	35	72.9	338	5	O01679	O01679 bombyx mori
17	35	72.9	381	16	Q8DSD4	Q8dsd4 streptococc
18	35	72.9	585	16	O66759	O66759 aquifex aeo
19	35	72.9	643	5	Q8IIV6	Q8iiv6 plasmodium
20	35	72.9	731	10	Q9AW75	Q9aw75 guillardia
21	35	72.9	767	10	Q8GXD0	Q8gxd0 arabidopsis
22	35	72.9	2655	10	Q9C8A3	Q9c8a3 arabidopsis
23	34	70.8	59	16	Q8YW77	Q8yw77 anabaena sp
24	34	70.8	89	9	Q9XJI1	Q9xjil bacteriopha
25	34	70.8	89	16	Q8XEH8	Q8xeh8 escherichia
26	34	70.8	115	16	Q8CV09	Q8cv09 oceanobacil
27	34	70.8	234	11	Q9D334	Q9d334 mus musculu
28	34	70.8	235	16	Q8XL88	Q8xl88 clostridium
29	34	70.8	245	4	O60573	O60573 homo sapien
30	34	70.8	245	11	O88503	O88503 mus musculu
31	34	70.8	245	11	Q8BMB3	Q8bmb3 mus musculu
32	34	70.8	276	16	Q9JTL7	Q9jtl7 neisseria m
33	34	70.8	285	16	O31436	O31436 bacillus su
34	34	70.8	287	16	Q98QY4	Q98qy4 mycoplasma
35	34	70.8	343	1	Q9HH38	Q9hh38 methanobact
36	34	70.8	432	2	Q8GCX1	Q8gcx1 enterococcu
37	34	70.8	521	5	Q9VPX7	Q9vpx7 drosophila
38	34	70.8	526	4	Q9Y3N2	Q9y3n2 homo sapien
39	34	70.8	547	16	Q8XE44	Q8xe44 escherichia
40	34	70.8	547	16	Q8FFQ4	Q8ffq4 escherichia
41	34	70.8	732	16	Q99XP7	Q99xp7 streptococc
42	34	70.8	732	16	Q8K5K3	Q8k5k3 streptococc
43	34	70.8	839	5	Q961M8	Q961m8 drosophila
44	34	70.8	954	4	Q8WXV8	Q8wxv8 homo sapien
45	34	70.8	957	4	Q9H0V3	Q9h0v3 homo sapien

ALIGNMENTS

RESULT 1
Q93CS6
ID Q93CS6 PRELIMINARY; PRT; 369 AA.
AC Q93CS6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Putative glycosyl transferase.
GN WBAT.
OS Shigella boydii.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=621;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21481970; PubMed=11598067;
RA Wang L., Qu W., Reeves P.R.;
RT "Sequence Analysis of Four Shigella boydii O-Antigen Loci: Implication
RT for Escherichia coli and Shigella Relationships.";
RL Infect. Immun. 69:6923-6930(2001).
DR EMBL; AF402314; AAL27338.1; -.
DR InterPro; IPR001296; Glyco_trans_1.
DR Pfam; PF00534; Glycos_transf_1; 1.
KW Transferase.
SQ SEQUENCE 369 AA; 42760 MW; 3E2CCE371A6BA0BB CRC64;

Query Match 83.3%; Score 40; DB 2; Length 369;
Best Local Similarity 77.8%; Pred. No. 23;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 KKWLIQLKK 9
              |||||:||:
Db        359 KKWLIKLKQ 367
```

RESULT 2
Q8MP31
ID Q8MP31 PRELIMINARY; PRT; 229 AA.
AC Q8MP31;

```
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AX4;
RA   Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Bear P.,
RA   Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA   Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT   "Sequence and Analysis of Chromosome 2 of Dictyostelium.";
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC117178; AAM44388.1; -.
DR   InterPro; IPR001064; Crystallin.
DR   PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   229 AA;  27735 MW;  FDC1E5C8B9647DD6 CRC64;

  Query Match             79.2%;  Score 38;  DB 5;  Length 229;
  Best Local Similarity   66.7%;  Pred. No. 34;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||| ::||
Db        119 KKWLTEIKK 127


RESULT 3
Q8Z1K3
ID   Q8Z1K3      PRELIMINARY;      PRT;   189 AA.
AC   Q8Z1K3;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Putative membrane protein.
GN   STY4561.
OS   Salmonella typhi.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Salmonella.
OX   NCBI_TaxID=601;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CT18;
RX   MEDLINE=21534947; PubMed=11677608;
RA   Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA   Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
RA   Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
RA   Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA   Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA   Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA   Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA   Whitehead S., Barrell B.G.;
RT   "Complete genome sequence of a multiple drug resistant Salmonella
RT   enterica serovar Typhi CT18.";
RL   Nature 413:848-852(2001).
DR   EMBL; AL627282; CAD09337.1; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   189 AA;  22077 MW;  C817745ABC69BE2A CRC64;

  Query Match             77.1%;  Score 37;  DB 16;  Length 189;
  Best Local Similarity   66.7%;  Pred. No. 44;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |:|||| |:
Db        101 KRWLIQAKR 109


RESULT 4
Q9JYL8
ID   Q9JYL8      PRELIMINARY;      PRT;   275 AA.
AC   Q9JYL8;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Phytoene synthase-related protein.
GN   NMB1521.
OS   Neisseria meningitidis (serogroup B).
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MC58 / Serogroup B;
RX   MEDLINE=20175755; PubMed=10710307;
RA   Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA   Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA   Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA   Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA   Mason T., Ciecko A., Parksey D.S., Blair E., Cittone H., Clark E.B.,
RA   Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
RA   Gill J., Scarlato V., Masignani V., Pizza M., Grandi G., Sun L.,
RA   Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT   "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT   MC58.";
RL   Science 287:1809-1815(2000).
DR   EMBL; AE002501; AAF41877.1; -.
DR   TIGR; NMB1521; -.
DR   InterPro; IPR002060; Squ/phyt_synthse.
DR   Pfam; PF00494; SQS_PSY; 1.
KW   Complete proteome.
SQ   SEQUENCE   275 AA;  30613 MW;  B8B2C2D77A549EA6 CRC64;

  Query Match             77.1%;  Score 37;  DB 16;  Length 275;
  Best Local Similarity   77.8%;  Pred. No. 62;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              | ||| |||
Db        263 KDWLIMLKK 271


RESULT 5
Q9EYA7
ID   Q9EYA7      PRELIMINARY;      PRT;    89 AA.
AC   Q9EYA7;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Kil protein.
GN   KIL.
OS   Escherichia coli O157:H7.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=83334;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O157:H7;
RX   MEDLINE=20564182; PubMed=11111050;
RA   Yokoyama K., Makino K., Kubota Y., Watanabe M., Kimura S.,
RA   Yutsudo C.H., Kurokawa K., Ishii K., Hattori M., Abe H., Iida T.,
RA   Yamamoto K., Hayashi T., Yasunaga T., Honda T., Sasakawa C.,
RA   Shinagawa H.;
RT   "Complete nucleotide sequence of the defective Sakai-VT1 prophage
RT   carrying the shiga toxin 1 genes of the enterohemorrhagic Escherichia
RT   coli O157:H7 derived from the Sakai outbreak.";
RL   Gene 258:127-139(2000).
DR   EMBL; AP000400; BAB19614.1; -.
SQ   SEQUENCE   89 AA;  10127 MW;  CD30904BF080D5FD CRC64;

  Query Match             75.0%;  Score 36;  DB 2;  Length 89;
  Best Local Similarity   55.6%;  Pred. No. 33;
  Matches    5;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy          1 KKWLIQLKK 9
              ||||: ||:
Db         74 KKWILMLKR 82


RESULT 6
Q9UJ05
ID   Q9UJ05      PRELIMINARY;      PRT;   278 AA.
AC   Q9UJ05;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   DJ486I3.1 (Novel protein) (Fragment).
GN   DJ486I3.1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Barlow K.;
RL   Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL050331; CAB55880.1; -.
FT   NON_TER       1       1
SQ   SEQUENCE   278 AA;  31677 MW;  0088AFE6B3BC3F6F CRC64;

  Query Match             75.0%;  Score 36;  DB 4;  Length 278;
  Best Local Similarity   87.5%;  Pred. No. 95;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLK 8
              |||| |||
Db         25 KKWLRQLK 32


RESULT 7
Q96BV3
ID   Q96BV3      PRELIMINARY;      PRT;   284 AA.
AC   Q96BV3;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Hypothetical protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Bladder;
RA   Strausberg R.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC015138; AAH15138.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   284 AA;  32370 MW;  BCD8F0A67699530A CRC64;

  Query Match             75.0%;  Score 36;  DB 4;  Length 284;
  Best Local Similarity   87.5%;  Pred. No. 97;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLK 8
              |||| |||
Db         31 KKWLRQLK 38


RESULT 8
Q8H3G5
ID   Q8H3G5      PRELIMINARY;      PRT;   459 AA.
AC   Q8H3G5;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   OSJNBb0062D12.24 protein.
GN   OSJNBB0062D12.24.
OS   Oryza sativa (japonica cultivar-group).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=39947;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Nipponbare;
RA   Sasaki T., Matsumoto T., Katayose Y.;
RT   "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 7, BAC
RT   clone:OSJNBb0062D12.";
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AP005127; BAC16708.1; -.
SQ   SEQUENCE   459 AA;  51464 MW;  5D66AF12D91CF90F CRC64;

  Query Match             75.0%;  Score 36;  DB 10;  Length 459;
  Best Local Similarity   75.0%;  Pred. No. 1.5e+02;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 KWLIQLKK 9
              |||:| ||
Db        322 KWLVQKKK 329


RESULT 9
O23488
ID   O23488      PRELIMINARY;      PRT;   578 AA.
AC   O23488;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein (LET1 like protein).
GN   AT4G16440.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98121113; PubMed=9461215;
RA   Bevan M., Bancroft I., Bent E., Love K., Goodman H., Dean C.,
RA   Bergkamp R., Dirkse W., van Staveren M., Stiekema W., Drost L.,
RA   Ridley P., Hudson S.A., Patel K., Murphy G., Piffanelli P., Wedler H.,
RA   Wedler E., Wambutt R., Weitzenegger T., Pohl T.M., Terryn N.,
RA   Gielen J., Villarroel R., De Clerck R., van Montagu M., Lecharny A.,
RA   Auborg S., Gy I., Kreis M., Lao N., Kavanagh T., Hempel S., Kotter P.,
RA   Entian K.D., Rieger M., Schaeffer M., Funk B., Mueller-Auer S.,
RA   Silvey M., James R., Montfort A., Pons A., Puigdomenech P., Douka A.,
RA   Voukelatou E., Milioni D., Hatzopoulos P., Piravandi E., Obermaier B.,
RA   Hilbert H., Duesterhoft A., Moores T., Jones J.D.G., Eneva T.,
RA   Palme K., Benes V., Rechman S., Ansorge W., Cooke R., Berger C.,
RA   Delseny M., Voet M., Volckaert G., Mewes H.W., Klosterman S.,
RA   Schueller C., Chalwatzis N.;
RT   "Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of
RT   Arabidopsis thaliana.";
RL   Nature 391:485-488(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z97341; CAB10420.1; -.
DR   EMBL; AL161544; CAB78686.1; -.
DR   InterPro; IPR006459; DUF_A_thal_1569.
DR   InterPro; IPR003149; Fe_hyd_SSU.
DR   Pfam; PF02256; Fe_hyd_SSU; 1.
DR   TIGRFAMs; TIGR01569; A_tha_TIGR01569; 1.
SQ   SEQUENCE   578 AA;  64340 MW;  212385A45DC3C09A CRC64;

  Query Match             75.0%;  Score 36;  DB 10;  Length 578;
  Best Local Similarity   87.5%;  Pred. No. 1.9e+02;
```

```
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 KWLIQLKK 9
              ||| ||||
Db         30 KWLDQLKK 37


RESULT 10
Q8I427
ID   Q8I427      PRELIMINARY;      PRT;   652 AA.
AC   Q8I427;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Cell differentiation protein rcd1, putative.
GN   PFE0375W.
OS   Plasmodium falciparum (isolate 3D7).
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=36329;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Devlin K., Baker S., Davies P., Mungal K., Berriman M., Pain A.,
RA   Hall N., Bowman S., Churcher C., Quail M., Barrell B.;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22255708; PubMed=12368867;
RA   Hall N., Pain A., Berriman M., Churcher C., Harris B., Harris D.,
RA   Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,
RA   Buckee C.O., Burrows C., Cherevach I., Chillingworth C.,
RA   Chillingworth T., Christodoulou Z., Clark L., Clark R., Corton C.,
RA   Cronin A., Davies R., Davis P., Dear P., Dearden F., Doggett J.,
RA   Feltwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N., Hance Z.,
RA   Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P.,
RA   Humphray S., Jagels K., James K.D., Johnson D., Kerhornou A.,
RA   Knights A., Konfortov B., Kyes S., Larke N., Lawson D., Lennard N.,
RA   Line A., Maddison M., Mclean J., Mooney P., Moule S., Murphy L.,
RA   Oliver K., Ormond D., Price C., Quail M.A., Rabbinowitsch E.,
RA   Rajandream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M.,
RA   Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K.,
RA   Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,
RA   Sulston J.E., Craig A., Newbold C., Barrell B.G;
RT   "Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13.";
RL   Nature 419:527-531(2002).
DR   EMBL; AL929351; CAD51441.1; -.
SQ   SEQUENCE   652 AA;  73423 MW;  90BBF62FBFB7FFFF CRC64;

  Query Match             75.0%;  Score 36;  DB 5;  Length 652;
  Best Local Similarity   85.7%;  Pred. No. 2.1e+02;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQL 7
              ||||:||
Db        482 KKWLLQL 488


RESULT 11
Q9N3E9
ID   Q9N3E9      PRELIMINARY;      PRT;  1566 AA.
AC   Q9N3E9;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical 180.6 kDa protein.
GN   Y54E10A.11.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RX   MEDLINE=99069613; PubMed=9851916;
RA   None;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology. The C. elegans Sequencing Consortium.";
RL   Science 282:2012-2018(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Wollam A., Becker M., Graves T., Hawkins M.;
RT   "The sequence of C. elegans cosmid Y54E10A.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Waterston R.;
RT   "Direct Submission.";
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC024810; AAF60772.2; -.
DR   WormPep; Y54E10A.11; CE29120.
DR   InterPro; IPR001841; Znf_ring.
DR   SMART; SM00184; RING; 1.
DR   PROSITE; PS50089; ZF_RING_2; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   1566 AA;  180607 MW;  F7D21D6EC9BA0F16 CRC64;

  Query Match             75.0%;  Score 36;  DB 5;  Length 1566;
  Best Local Similarity   85.7%;  Pred. No. 4.7e+02;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQL 7
              ||||:||
Db       1477 KKWLLQL 1483


RESULT 12
Q8EI76
ID   Q8EI76      PRELIMINARY;      PRT;   136 AA.
AC   Q8EI76;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein.
GN   SO0971.
OS   Shewanella oneidensis.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC   Alteromonadaceae; Shewanella.
OX   NCBI_TaxID=70863;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MR-1;
RX   MEDLINE=22297686; PubMed=12368813;
RA   Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA   Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,
RA   Meyer T., Tsapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,
RA   DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
RA   Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
RA   Vamathevan J., Weidman J., Impraim M., Lee K., Berry K., Lee C.,
RA   Mueller J., Khouri H., Gill J., Utterback T.R., McDonald L.A.,
RA   Feldblyum T.V., Smith H.O., Venter J.C., Nealson K.H., Fraser C.M.;
RT   "Genome sequence of the dissimilatory metal ion-reducing bacterium
RT   Shewanella oneidensis.";
RL   Nat. Biotechnol. 20:1118-1123(2002).
DR   EMBL; AE015542; AAN54045.1; -.
DR   TIGR; SO0971; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   136 AA;  15529 MW;  420710910979605C CRC64;

  Query Match             72.9%;  Score 35;  DB 16;  Length 136;
  Best Local Similarity   85.7%;  Pred. No. 75;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 KWLIQLK 8
              :||||||
Db         69 RWLIQLK 75
```

```
RESULT 13
O65062
ID   O65062      PRELIMINARY;      PRT;   198 AA.
AC   O65062;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Rac-like GTP binding protein.
GN   SB30.
OS   Picea mariana (Black spruce).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Picea.
OX   NCBI_TaxID=3335;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98278823; PubMed=9611216;
RA   Perry D.J., Bousquet J.;
RT   "Sequence-tagged-site (STS) markers of arbitrary genes. Development,
RT   characterization and analysis of linkage in black spruce.";
RL   Genetics 149:1089-1098(1998).
CC   -!- SIMILARITY: BELONGS TO THE SMALL GTPASE SUPERFAMILY. RHO FAMILY.
DR   EMBL; AF051223; AAC32124.1; -.
DR   HSSP; P21181; 1AM4.
DR   InterPro; IPR003578; GTPase_Rho.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00174; RHO; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
KW   GTP-binding; Lipoprotein; Prenylation.
SQ   SEQUENCE   198 AA;  21997 MW;  FA72F6EBB1D15E45 CRC64;

  Query Match             72.9%;  Score 35;  DB 10;  Length 198;
  Best Local Similarity   62.5%;  Pred. No. 1.1e-02;
  Matches    5;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLK 8
              ||||::|:
Db         98 KKWLLELR 105


RESULT 14
Q9CNC3
ID   Q9CNC3      PRELIMINARY;      PRT;   261 AA.
AC   Q9CNC3;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Hypothetical protein PM0509.
GN   PM0509.
OS   Pasteurella multocida.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Pasteurella.
OX   NCBI_TaxID=747;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Pm70;
RX   MEDLINE=21145866; PubMed=11248100;
RA   May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT   "Complete genomic sequence of Pasteurella multocida Pm70.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
DR   EMBL; AE006086; AAK02593.1; -.
DR   InterPro; IPR002654; Glyco_trans_25.
DR   Pfam; PF01755; Glyco_transf_25; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   261 AA;  30190 MW;  6D71C37C615EEF44 CRC64;

  Query Match             72.9%;  Score 35;  DB 16;  Length 261;
  Best Local Similarity   66.7%;  Pred. No. 1.4e-02;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              :|||| |:|
Db          2 QKWLISLEK 10


RESULT 15
Q8HIS8
ID   Q8HIS8      PRELIMINARY;      PRT;   294 AA.
AC   Q8HIS8;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Orf294.
GN   ORF294.
OS   Monosiga brevicollis.
OG   Mitochondrion.
OC   Eukaryota; Choanoflagellida; Codonosigidae; Monosiga.
OX   NCBI_TaxID=81824;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20377911; PubMed=10916154;
RA   Bullerwell C.E., Burger G., Lang B.F.;
RT   "A novel motif for identifying rps3 homologs in fungal mitochondrial
RT   genomes.";
RL   Trends Biochem. Sci. 25:363-365(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Lang B.F., O'Kelly C.J., Nerad T.A., Gray M.W., Burger G.;
RT   "The closest unicellular relatives of animals.";
RL   Curr. Biol. 0:0-0(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Lang F.B., Bullerwell C.;
RL   Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Burger G., Lang B.F.;
RL   Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF538053; AAN28356.1; -.
KW   Mitochondrion.
SQ   SEQUENCE   294 AA;  34521 MW;  6CD1601F13C6D9D6 CRC64;

  Query Match             72.9%;  Score 35;  DB 8;  Length 294;
  Best Local Similarity   75.0%;  Pred. No. 1.5e+02;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLK 8
              |||:| ||
Db         30 KKWIITLK 37


Search completed: October 14, 2003, 14:36:55
Job time : 32.8125 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 14:33:19 ; Search time 38.625 Seconds
(without alignments)
36.985 Million cell updates/sec

Title: US-09-881-490-163
Perfect score: 48
Sequence: 1 KKLQILWKK 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	48	100.0	9	17	AAW04048	Antifungal peptide
2	48	100.0	9	18	AAW44562	Anti-fungal peptid
3	48	100.0	9	20	AAY00539	Antifungal peptide
4	48	100.0	9	22	AAB65463	Anti-fungal peptid
5	43	89.6	8	18	AAW44604	Anti-fungal peptid
6	43	89.6	8	20	AAY00581	Antifungal peptide
7	43	89.6	8	22	AAB65505	Anti-fungal peptid
8	43	89.6	9	18	AAW44606	Anti-fungal peptid
9	43	89.6	9	20	AAY00583	Antifungal peptide
10	43	89.6	9	22	AAB65507	Anti-fungal peptid
11	38	79.2	128	22	AAO00543	Human polypeptide
12	38	79.2	325	23	ABG65564	Human breast speci
13	38	79.2	369	22	ABG15487	Novel human diagno
14	38	79.2	414	22	ABG15486	Novel human diagno
15	36	75.0	547	23	ABB05745	Mouse zcytor17 pro
16	36	75.0	662	23	ABB05742	Mouse zcytor17 pro
17	36	75.0	716	24	ABP54370	Mouse mNR10Balb-c
18	36	75.0	716	24	ABP54371	Mouse mNR10c57BL-6
19	36	75.0	726	23	AAE24037	Mouse haematopoiet
20	36	75.0	2086	22	ABB64128	Drosophila melanog
21	35	72.9	55	21	AAG01005	Human secreted pro
22	35	72.9	145	22	AAO01497	Human polypeptide
23	35	72.9	303	20	AAW88847	Polypeptide fragme
24	35	72.9	303	22	ABB50865	Human secreted pro
25	35	72.9	400	20	AAW88851	Polypeptide fragme
26	35	72.9	400	22	ABB50869	Human secreted pro
27	35	72.9	492	22	ABG03810	Novel human diagno
28	35	72.9	492	22	ABG14087	Novel human diagno
29	35	72.9	738	20	AAW95332	Human phospholipas
30	35	72.9	740	21	AAY58878	Human peptidase NA
31	35	72.9	740	22	AAM24008	Human EST encoded
32	35	72.9	740	22	AAB88354	Human membrane or
33	35	72.9	795	21	AAY95968	Human TANGO 243.
34	35	72.9	795	22	AAB93162	Human protein sequ
35	35	72.9	795	22	AAB68514	Human GFP-binding
36	35	72.9	855	21	AAY95969	Human TANGO 243 (a
37	34	70.8	18	21	AAB22991	Amphipathic cation
38	34	70.8	20	19	AAW65895	Amphipathic peptid
39	34	70.8	20	21	AAY49957	Amphipathic peptid
40	34	70.8	45	21	AAG07026	Arabidopsis thalia
41	34	70.8	49	22	ABG58565	Human liver peptid
42	34	70.8	49	22	ABB43169	Peptide #10675 enc
43	34	70.8	49	22	ABB26273	Protein #8272 enco
44	34	70.8	49	22	AAM64078	Human brain expres
45	34	70.8	49	22	AAM76899	Human bone marrow

ALIGNMENTS

RESULT 1
AAW04048
ID AAW04048 standard; peptide; 9 AA.
XX
AC AAW04048;
XX
DT 04-NOV-1996 (first entry)
XX
DE Antifungal peptide XMP.332.
XX
KW Antifungal peptide; inhibitor; Domain III; polymorphonuclear leukocyte;
KW bactericidal/permeability-increasing protein; BPI; mammalian; PMN; fungi;
KW neutrophil; replication inhibitor; fungal infection; Aspergillus;
KW Cryptococcus; Candida; C.albicans; C.galabrat; C.krusei; C.lusitaniae;
KW C.parapsilosis; C.tropicalis; therapy.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 1..9
FT /note= "D-form residues"
FT Modified-site 9
FT /note= "amidated"
XX
PN WO9608509-A1.
XX
PD 21-MAR-1996.
XX
PF 20-JUL-1995; 95WO-US09262.
XX
PR 13-JAN-1995; 95US-0372105.
PR 15-SEP-1994; 94US-0306473.

XX
PA (XOMA) XOMA CORP.
XX
PI Fadem MB, Lim E, Little RG;
XX
DR WPI; 1996-179900/18.
XX
PT Antifungal peptide(s) derived from Domain III of BPI protein - used
PT in vitro for killing or inhibiting replication of fungi, esp.
PT Candida species
XX
PS Claim 9; Page 155; 199pp; English.
XX
CC AAW04000-W04160 represent antifungal peptides. These sequences are
CC based on (or derived from) Domain III of the
CC bactericidal/permeability-increasing protein (BPI). BPI is a protein
CC that can be isolated from the granules of mammalian polymorphonuclear
CC leukocytes (PMNs or neutrophils). These antifungal peptides can be used
CC for killing, or inhibiting replication of, fungi in vitro. These
CC sequences can also be used for treatment of a fungal infection involving
CC fungi from the species Candida, Aspergillus and Cryptococcus. The
CC sequences are especially useful for treating C.albicans, C.galabrat,
CC C.krusei, C.lusitaniae, C.parapsilosis and C.tropicalis infections.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 48; DB 17; Length 9;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKLQILWKK 9
|||||||||
Db 1 KKLQILWKK 9

RESULT 2
AAW44562
ID AAW44562 standard; peptide; 9 AA.
XX
AC AAW44562;
XX
DT 27-APR-1998 (first entry)
XX
DE Anti-fungal peptide #163 based on BPI protein (residues 142-169).
XX
KW Anti-fungal peptide; bactericidal-permeability-increasing protein; BPI;
KW polymorphonuclear leukocyte; fungicide.
XX
OS Synthetic.
OS Mammalia.
XX
FH Key Location/Qualifiers
FT Misc-difference 1..9
FT /note= "D-form residues"
FT Modified-site 9
FT /note= "C-terminal amide"
XX
PN WO9704008-A1.
XX
PD 06-FEB-1997.
XX
PF 21-MAR-1996; 96WO-US03845.
XX
PR 20-JUL-1995; 95US-0504841.
XX
PA (XOMA) XOMA CORP.
XX
PI Fadem MB, Lim E, Little RG;
XX
DR WPI; 1997-132578/12.
XX
PT Anti-fungal peptide(s) derived from or based on domain III of
PT bactericidal-permeability-increasing protein - are used in vitro or
PT in vivo as a fungicides
XX
PS Claim 1; Page 192; 230pp; English.
XX
CC This is a specifically claimed anti-fungal peptide which is based on
CC domain III (amino acids 142-160) of bactericidal-permeability-increasing
CC protein (BPI), isolated from the granules of mammalian polymorphonuclear
CC leukocytes. It is used in compositions with diluents, carriers or
CC adjuvants to treat fungal infections in patients. It may also be used in
CC vitro to kill or inhibit the replication of fungi, such as in
CC decontaminating fluids and sterilising medical and implant devices.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 48; DB 18; Length 9;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKLQILWKK 9
|||||||||
Db 1 KKLQILWKK 9

RESULT 3
AAY00539
ID AAY00539 standard; Peptide; 9 AA.
XX
AC AAY00539;
XX
DT 07-MAY-1999 (first entry)
XX
DE Antifungal peptide XMP.332.
XX
KW Antifungal; BPI; bactericidal/permeability increasing protein;
KW Candida infection.
XX
OS Synthetic.
XX
PN US5858974-A.
XX
PD 12-JAN-1999.
XX
PF 21-MAR-1996; 96US-0621259.
XX
PR 21-MAR-1996; 96US-0621259.
PR 20-JUL-1995; 95US-0504841.
XX
PA (XOMA) XOMA CORP.
XX
PI Fadem MB, Lim E, Little RG;
XX
DR WPI: 1999-119956/10.
XX
PT Antifungal peptides - comprising part of bactericidal or
PT permeability-increasing protein sequence or related sequence
XX
PS Disclosure; Columns 167-168; 132pp; English.
XX
CC New peptides are provided which are based on Domain III (amino acids
CC 142-169) of human bactericidal/permeability-increasing protein (BPI).
CC The peptides all have a C-terminal amide. More particularly, the Claims
CC relate to: (1) a peptide that has an amino acid sequence of human BPI
CC from position 148 to position 161 (KSKVGWLIQLFHKK) and variants of the
CC sequence having antifungal activity; and (2) an antifungal peptide
CC having 6-14 amino acids comprising (a) a core sequence selected from
CC LIQL, IQLF, WLIQL, LIQLF and WLIQLF and (b) one or more cationic amino
CC acids selected from K, R, H, Orn (ornithine) and Dab (diaminobutyric
CC acid) at the N and/or C terminus of the core sequence. The new peptides
CC are used for killing or inhibiting replication of fungi in vitro; and
CC for treating fungal infections in vivo, in particular infections of
CC Candida, Aspergillus or Cryptococcus spp., especially C. albicans, C.
CC krusei, C. lusitaniae, C. parapsilosis or C. tropicalis. The peptide
CC can be administered topically, intravenously, orally or as an aerosol,

```
CC   optionally together with a non-peptide antifungal agent.
XX
SQ   Sequence   9 AA;

  Query Match            100.0%;  Score 48;  DB 20;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 9.3e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLQILWKK 9
              |||||||||
Db          1 KKLQILWKK 9


RESULT 4
AAB65463
ID   AAB65463 standard; Peptide; 9 AA.
XX
AC   AAB65463;
XX
DT   27-MAR-2001  (first entry)
XX
DE   Anti-fungal peptide XMP.332.
XX
KW   Human; BPI; antifungal; polymorphonuclear leukocyte; neutrophil;
KW   bactericidal/permeability-increasing protein; bactericidal;
KW   fungal infection.
XX
OS   Homo sapiens.
XX
PN   US6156730-A.
XX
PD   05-DEC-2000.
XX
PF   08-JAN-1999;   99US-0227659.
XX
PR   21-MAR-1996;   96US-0621259.
PR   12-MAR-1993;   93US-0030644.
PR   15-JUL-1993;   93US-0093202.
PR   14-JAN-1994;   94US-0183222.
PR   11-MAR-1994;   94US-0209762.
PR   11-JUL-1994;   94US-0273540.
PR   15-SEP-1994;   94US-0306473.
PR   13-JAN-1995;   95US-0372105.
PR   20-JUL-1995;   95US-0504841.
XX
PA   (XOMA ) XOMA CORP.
XX
PI   Lim E,  Fadem MB,  Little RG;
XX
DR   WPI; 2001-090160/10.
XX
PT   Novel anti-fungal peptides derived from domain III of
PT   bactericidal/permeability-increasing protein useful for killing or
PT   inhibiting replication of fungi and for treating fungal infections -
XX
PS   Example 2; Columns 171-172; 134pp; English.
XX
CC   The present invention relates to antifungal peptides (see
CC   AAB65301-B65550) derived from Domain III (amino acids 142-169) of
CC   bactericidal/permeability-increasing protein (BPI). The present sequence
CC   is one such antifungal peptide. BPI is a protein isolated from the
CC   granules of mammalian polymorphonuclear leukocytes (PMNs or
CC   neutrophils). BPI has potent bactericidal activity against a broad range
CC   of gram-negative bacteria. The peptides of the present invention are
CC   useful for killing or inhibiting replication of fungi, and treating
CC   infections caused by fungus belonging to Candida, Aspergillus,
CC   Cryptococcus species such as C.albicans, C.glabrata, C.krusei,
CC   C.lusitaniae, C.parapsilosis and C.tropicalis.
XX
SQ   Sequence   9 AA;

  Query Match            100.0%;  Score 48;  DB 22;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 9.3e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLQILWKK 9
              |||||||||
Db          1 KKLQILWKK 9


RESULT 5
AAW44604
ID   AAW44604 standard; peptide; 8 AA.
XX
AC   AAW44604;
XX
DT   27-APR-1998  (first entry)
XX
DE   Anti-fungal peptide #205 based on BPI protein (residues 142-169).
XX
KW   Anti-fungal peptide; bactericidal-permeability-increasing protein; BPI;
KW   polymorphonuclear leukocyte; fungicide.
XX
OS   Synthetic.
OS   Mammalia.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1..8
FT                   /note= "D-form residues"
FT   Modified-site   8
FT                   /note= "C-terminal amide"
XX
PN   WO9704008-A1.
XX
PD   06-FEB-1997.
XX
PF   21-MAR-1996;   96WO-US03845.
XX
PR   20-JUL-1995;   95US-0504841.
XX
PA   (XOMA ) XOMA CORP.
XX
PI   Fadem MB,  Lim E,  Little RG;
XX
DR   WPI; 1997-132578/12.
XX
PT   Anti-fungal peptide(s) derived from or based on domain III of
PT   bactericidal-permeability-increasing protein - are used in vitro or
PT   in vivo as a fungicides
XX
PS   Claim 1; -pp; 230pp; English.
XX
CC   This is a specifically claimed anti-fungal peptide which is based on
CC   domain III (amino acids 142-160) of bactericidal-permeability-increasing
CC   protein (BPI), isolated from the granules of mammalian polymorphonuclear
CC   leukocytes. It is used in compositions with diluents, carriers or
CC   adjuvants to treat fungal infections in patients. It may also be used in
CC   vitro to kill or inhibit the replication of fungi, such as in
CC   decontaminating fluids and sterilising medical and implant devices.
XX
SQ   Sequence   8 AA;

  Query Match            89.6%;  Score 43;  DB 18;  Length 8;
  Best Local Similarity  100.0%;  Pred. No. 9.3e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 KLQILWKK 9
              ||||||||
Db          1 KLQILWKK 8


RESULT 6
AAY00581
ID   AAY00581 standard; Peptide; 8 AA.
XX
AC   AAY00581;
```

XX
DT 07-MAY-1999 (first entry)
XX
DE Antifungal peptide XMP.374.
XX
KW Antifungal; BPI; bactericidal/permeability increasing protein;
KW Candida infection.
XX
OS Synthetic.
XX
PN US5858974-A.
XX
PD 12-JAN-1999.
XX
PF 21-MAR-1996; 96US-0621259.
XX
PR 21-MAR-1996; 96US-0621259.
PR 20-JUL-1995; 95US-0504841.
XX
PA (XOMA) XOMA CORP.
XX
PI Fadem MB, Lim E, Little RG;
XX
DR WPI; 1999-119956/10.
XX
PT Antifungal peptides - comprising part of bactericidal or
PT permeability-increasing protein sequence or related sequence
XX
PS Claim 2; Columns 197-198; 132pp; English.
XX
CC New peptides are provided which are based on Domain III (amino acids
CC 142-169) of human bactericidal/permeability-increasing protein (BPI).
CC The peptides all have a C-terminal amide. More particularly, the Claims
CC relate to: (1) a peptide that has an amino acid sequence of human BPI
CC from position 148 to position 161 (KSKVGWLIQLFHKK) and variants of the
CC sequence having antifungal activity; and (2) an antifungal peptide
CC having 6-14 amino acids comprising (a) a core sequence selected from
CC LIQL, IQLF, WLIQL, LIQLF and WLIQLF and (b) one or more cationic amino
CC acids selected from K, R, H, Orn (ornithine) and Dab (diaminobutyric
CC acid) at the N and/or C terminus of the core sequence. The new peptides
CC are used for killing or inhibiting replication of fungi in vitro; and
CC for treating fungal infections in vivo, in particular infections of
CC Candida, Aspergillus or Cryptococcus spp., especially C. albicans, C.
CC krusei, C. lusitaniae, C. parapsilosis or C. tropicalis. The peptide
CC can be administered topically, intravenously, orally or as an aerosol,
CC optionally together with a non-peptide antifungal agent.
XX
SQ Sequence 8 AA;

Query Match 89.6%; Score 43; DB 20; Length 8;
Best Local Similarity 100.0%; Pred. No. 9.3e-05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 KLQILWKK 9
||||||||
Db 1 KLQILWKK 8

RESULT 7
AAB65505
ID AAB65505 standard; Peptide; 8 AA.
XX
AC AAB65505;
XX
DT 27-MAR-2001 (first entry)
XX
DE Anti-fungal peptide XMP.374.
XX
KW Human; BPI; antifungal; polymorphonuclear leukocyte; neutrophil;
KW bactericidal/permeability-increasing protein; bactericidal;
KW fungal infection.
XX
OS Homo sapiens.

XX
PN US6156730-A.
XX
PD 05-DEC-2000.
XX
PF 08-JAN-1999; 99US-0227659.
XX
PR 21-MAR-1996; 96US-0621259.
PR 12-MAR-1993; 93US-0030644.
PR 15-JUL-1993; 93US-0093202.
PR 14-JAN-1994; 94US-0183222.
PR 11-MAR-1994; 94US-0209762.
PR 11-JUL-1994; 94US-0273540.
PR 15-SEP-1994; 94US-0306473.
PR 13-JAN-1995; 95US-0372105.
PR 20-JUL-1995; 95US-0504841.
XX
PA (XOMA) XOMA CORP.
XX
PI Lim E, Fadem MB, Little RG;
XX
DR WPI; 2001-090160/10.
XX
PT Novel anti-fungal peptides derived from domain III of
PT bactericidal/permeability-increasing protein useful for killing or
PT inhibiting replication of fungi and for treating fungal infections -
XX
PS Example 2; Columns 203-204; 134pp; English.
XX
CC The present invention relates to antifungal peptides (see
CC AAB65301-B65550) derived from Domain III (amino acids 142-169) of
CC bactericidal/permeability-increasing protein (BPI). The present sequence
CC is one such antifungal peptide. BPI is a protein isolated from the
CC granules of mammalian polymorphonuclear leukocytes (PMNs or
CC neutrophils). BPI has potent bactericidal activity against a broad range
CC of gram-negative bacteria. The peptides of the present invention are
CC useful for killing or inhibiting replication of fungi, and treating
CC infections caused by fungus belonging to Candida, Aspergillus,
CC Cryptococcus species such as C.albicans, C.glabrata, C.krusei,
CC C.lusitaniae, C.parapsilosis and C.tropicalis.
XX
SQ Sequence 8 AA;

Query Match 89.6%; Score 43; DB 22; Length 8;
Best Local Similarity 100.0%; Pred. No. 9.3e-05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 KLQILWKK 9
||||||||
Db 1 KLQILWKK 8

RESULT 8
AAW44606
ID AAW44606 standard; peptide; 9 AA.
XX
AC AAW44606;
XX
DT 27-APR-1998 (first entry)
XX
DE Anti-fungal peptide #207 based on BPI protein (residues 142-169).
XX
KW Anti-fungal peptide; bactericidal-permeability-increasing protein; BPI;
KW polymorphonuclear leukocyte; fungicide.
XX
OS Synthetic.
OS Mammalia.
XX
FH Key Location/Qualifiers
FT Misc-difference 1..9
FT /note= "D-form residues"
FT Modified-site 9
FT /note= "C-terminal amide"

XX
PN WO9704008-A1.
XX
PD 06-FEB-1997.
XX
PF 21-MAR-1996; 96WO-US03845.
XX
PR 20-JUL-1995; 95US-0504841.
XX
PA (XOMA) XOMA CORP.
XX
PI Fadem MB, Lim E, Little RG;
XX
DR WPI; 1997-132578/12.
XX
PT Anti-fungal peptide(s) derived from or based on domain III of
PT bactericidal-permeability-increasing protein - are used in vitro or
PT in vivo as a fungicides
XX
PS Claim 1; -PP; 230pp; English.
XX
CC This is a specifically claimed anti-fungal peptide which is based on
CC domain III (amino acids 142-160) of bactericidal-permeability-increasing
CC protein (BPI), isolated from the granules of mammalian polymorphonuclear
CC leukocytes. It is used in compositions with diluents, carriers or
CC adjuvants to treat fungal infections in patients. It may also be used in
CC vitro to kill or inhibit the replication of fungi, such as in
CC decontaminating fluids and sterilising medical and implant devices.
XX
SQ Sequence 9 AA;

Query Match 89.6%; Score 43; DB 18; Length 9;
Best Local Similarity 88.9%; Pred. No. 9.3e+05;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KKLQILWKK 9
||||| ||
Db 1 KKLQIAWKK 9

RESULT 9
AAY00583
ID AAY00583 standard; Peptide; 9 AA.
XX
AC AAY00583;
XX
DT 07-MAY-1999 (first entry)
XX
DE Antifungal peptide XMP.376.
XX
KW Antifungal; BPI; bactericidal/permeability increasing protein;
KW Candida infection.
XX
OS Synthetic.
XX
PN US5858974-A.
XX
PD 12-JAN-1999.
XX
PF 21-MAR-1996; 96US-0621259.
XX
PR 21-MAR-1996; 96US-0621259.
PR 20-JUL-1995; 95US-0504841.
XX
PA (XOMA) XOMA CORP.
XX
PI Fadem MB, Lim E, Little RG;
XX
DR WPI; 1999-119956/10.
XX
PT Antifungal peptides - comprising part of bactericidal or
PT permeability-increasing protein sequence or related sequence
XX
PS Claim 2; Columns 199-200; 132pp; English.
XX
CC New peptides are provided which are based on Domain III (amino acids
CC 142-169) of human bactericidal/permeability-increasing protein (BPI).
CC The peptides all have a C-terminal amide. More particularly, the Claims
CC relate to: (1) a peptide that has an amino acid sequence of human BPI
CC from position 148 to position 161 (KSKVGWLIQLFHKK) and variants of the
CC sequence having antifungal activity; and (2) an antifungal peptide
CC having 6-14 amino acids comprising (a) a core sequence selected from
CC LIQL, IQLF, WLIQL, LIQLF and WLIQLF and (b) one or more cationic amino
CC acids selected from K, R, H, Orn (ornithine) and Dab (diaminobutyric
CC acid) at the N and/or C terminus of the core sequence. The new peptides
CC are used for killing or inhibiting replication of fungi in vitro; and
CC for treating fungal infections in vivo, in particular infections of
CC Candida, Aspergillus or Cryptococcus spp., especially C. albicans, C.
CC krusei, C. lusitaniae, C. parapsilosis or C. tropicalis. The peptide
CC can be administered topically, intravenously, orally or as an aerosol,
CC optionally together with a non-peptide antifungal agent.
XX
SQ Sequence 9 AA;

Query Match 89.6%; Score 43; DB 20; Length 9;
Best Local Similarity 88.9%; Pred. No. 9.3e+05;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KKLQILWKK 9
||||| ||
Db 1 KKLQIAWKK 9

RESULT 10
AAB65507
ID AAB65507 standard; Peptide; 9 AA.
XX
AC AAB65507;
XX
DT 27-MAR-2001 (first entry)
XX
DE Anti-fungal peptide XMP.376.
XX
KW Human; BPI; antifungal; polymorphonuclear leukocyte; neutrophil;
KW bactericidal/permeability-increasing protein; bactericidal;
KW fungal infection.
XX
OS Homo sapiens.
XX
PN US6156730-A.
XX
PD 05-DEC-2000.
XX
PF 08-JAN-1999; 99US-0227659.
XX
PR 21-MAR-1996; 96US-0621259.
PR 12-MAR-1993; 93US-0030644.
PR 15-JUL-1993; 93US-0093202.
PR 14-JAN-1994; 94US-0183222.
PR 11-MAR-1994; 94US-0209762.
PR 11-JUL-1994; 94US-0273540.
PR 15-SEP-1994; 94US-0306473.
PR 13-JAN-1995; 95US-0372105.
PR 20-JUL-1995; 95US-0504841.
XX
PA (XOMA) XOMA CORP.
XX
PI Lim E, Fadem MB, Little RG;
XX
DR WPI; 2001-090160/10.
XX
PT Novel anti-fungal peptides derived from domain III of
PT bactericidal/permeability-increasing protein useful for killing or
PT inhibiting replication of fungi and for treating fungal infections -
XX
PS Example 2; Columns 205-206; 134pp; English.

```
XX
CC   The present invention relates to antifungal peptides (see
CC   AAB65301-B65550) derived from Domain III (amino acids 142-169) of
CC   bactericidal/permeability-increasing protein (BPI). The present sequence
CC   is one such antifungal peptide. BPI is a protein isolated from the
CC   granules of mammalian polymorphonuclear leukocytes (PMNs or
CC   neutrophils). BPI has potent bactericidal activity against a broad range
CC   of gram-negative bacteria. The peptides of the present invention are
CC   useful for killing or inhibiting replication of fungi, and treating
CC   infections caused by fungus belonging to Candida, Aspergillus,
CC   Cryptococcus species such as C.albicans, C.glabrata, C.krusei,
CC   C.lusitaniae, C.parapsilosis and C.tropicalis.
XX
SQ   Sequence   9 AA;

  Query Match            89.6%;  Score 43;  DB 22;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 9.3e+05;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKLQILWKK 9
              ||||| ||:
Db          1 KKLQIAWKK 9


RESULT 11
AAO00543
ID   AAO00543 standard; Protein; 128 AA.
XX
AC   AAO00543;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human polypeptide SEQ ID NO 14435.
XX
KW   Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW   vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW   tissue growth factor; immunomodulatory; cancer; leukaemia;
KW   nervous system disorders; arthritis; inflammation.
XX
OS   Homo sapiens.
XX
PN   WO200164835-A2.
XX
PD   07-SEP-2001.
XX
PF   26-FEB-2001; 2001WO-US04927.
XX
PR   28-FEB-2000; 2000US-0515126.
PR   18-MAY-2000; 2000US-0577409.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Drmanac RT;
XX
DR   WPI; 2001-514838/56.
DR   N-PSDB; AAI80474.
XX
PT   Isolated nucleic acids and polypeptides, useful for preventing
PT   diagnosing and treating e.g. leukaemia, inflammation and immune
PT   disorders -
XX
PS   Claim 20; SEQ ID NO 14435; 1399pp + Sequence Listing; English.
XX
CC   The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC   the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
CC   cytokine, cell proliferation or cell differentiation or which may induce
CC   production of other cytokines in other cell populations. The
CC   polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC   peptide therapy. The polypeptides have various cytokine-like activities,
CC   e.g. stem cell growth factor activity, haematopoiesis regulating
CC   activity, tissue growth factor activity, immunomodulatory activity and
CC   activin/inhibin activity and may be useful in the diagnosis and/or
CC   treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC   inflammation.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   128 AA;

  Query Match            79.2%;  Score 38;  DB 22;  Length 128;
  Best Local Similarity   66.7%;  Pred. No. 54;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKLQILWKK 9
              | :|:|||| 
Db         58 KMMQLLWKK 66


RESULT 12
ABG65564
ID   ABG65564 standard; Protein; 325 AA.
XX
AC   ABG65564;
XX
DT   23-AUG-2002  (first entry)
XX
DE   Human breast specific polypeptide, BSP #34.
XX
KW   Human; breast specific protein; BSP; breast cancer; vaccine;
KW   mammary tumour; cytostatic; gene therapy; non-cancerous breast disorder.
XX
OS   Homo sapiens.
XX
PN   WO200236807-A2.
XX
PD   10-MAY-2002.
XX
PF   29-OCT-2001; 2001WO-US46888.
XX
PR   27-OCT-2000; 2000US-243802P.
XX
PA   (DIAD-) DIADEXUS INC.
XX
PI   Sun Y,  Recipon H,  Salceda S,  Liu C,  Turner LR;
XX
DR   WPI; 2002-463415/49.
XX
PT   New breast-specific nucleic acids and polypeptides, useful for
PT   identifying, diagnosing, monitoring, staging, imaging, and treating
PT   breast cancer and non-cancerous disease states in breast tissues -
XX
PS   Claim 11; Page 231-232; 281pp; English.
XX
CC   The invention relates to breast specific nucleic acids (BSNA) and
CC   breast-specific polypeptides (BSP). Also included are a method for
CC   determining the BSNA in a sample, a vector comprising a BSNA, a host cell
CC   comprising the vector, a method for producing a polypeptide encoded
CC   by a BSNA, an anti-BSP antibody and a method for determining the presence
CC   of a BSP in a sample. The breast-specific nucleic acids, polypeptides and
CC   compositions comprising them are useful for identifying, diagnosing,
CC   monitoring, staging, imaging, and treating breast cancer, mammary tumour
CC   and non-cancerous disease states in breast tissue; for identifying breast
CC   tissue; for monitoring, identifying and/or designing agonists and
CC   antagonists of the polypeptides; in gene therapy; in producing transgenic
CC   animals and cells; for producing engineered breast tissue for treatment
CC   and research; and as elements in an array or computer program for pattern
CC   recognition of breast disorders. The nucleic acids may be used as
CC   hybridisation probes to detect, characterise and quantify hybridising
CC   nucleic acids in, and isolate hybridising nucleic acids from, both
CC   genomic and transcript-derived nucleic acid samples. The BSP
CC   protein may be used in a vaccine composition for raising an immune
CC   response against breast cancer. The present sequence is BSP protein of
CC   the invention.
XX
SQ   Sequence   325 AA;
```

```
  Query Match           79.2%;  Score 38;  DB 23;  Length 325;
  Best Local Similarity  75.0%;  Pred. No. 1.3e+02;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 KLQILWKK 9
              |||::|||
Db         78 KLQLVWKK 85


RESULT 13
ABG15487
ID   ABG15487 standard; Protein; 369 AA.
XX
AC   ABG15487;
XX
DT   18-FEB-2002  (first entry)
XX
DE   Novel human diagnostic protein #15478.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   N-PSDB; AAS79674.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity  -
XX
PS   Claim 20; SEQ ID No 45846; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. ABG00010-ABG30377 represent novel human
CC   diagnostic amino acid sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   369 AA;

  Query Match           79.2%;  Score 38;  DB 22;  Length 369;
  Best Local Similarity  66.7%;  Pred. No. 1.5e+02;
  Matches    6;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLQILWKK 9
              |:::|||||
Db        172 KQMRILWKK 180


RESULT 14
ABG15486
ID   ABG15486 standard; Protein; 414 AA.
XX
AC   ABG15486;
XX
DT   18-FEB-2002  (first entry)
XX
DE   Novel human diagnostic protein #15477.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   N-PSDB; AAS79673.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity  -
XX
PS   Claim 20; SEQ ID No 45845; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. ABG00010-ABG30377 represent novel human
CC   diagnostic amino acid sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   414 AA;

  Query Match           79.2%;  Score 38;  DB 22;  Length 414;
  Best Local Similarity  66.7%;  Pred. No. 1.6e+02;
  Matches    6;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 KKLQILWKK 9
              |:::|||||
Db        310 KQMRILWKK 318
```

RESULT 15
ABB05745
ID ABB05745 standard; Protein; 547 AA.
XX
AC ABB05745;
XX
DT 01-MAY-2002 (first entry)
XX
DE Mouse zcytor17 protein SEQ ID NO:93.
XX
KW Zcytor17; chromosome 5; 5q11; cytokine receptor; immunomodulatory;
KW antiinflammatory; antiviral; antirheumatic; antiarthritic; cytostatic;
KW muscular; lymphoid; immune; inflammatory; spleenic; blood; bone;
KW infection; immunosuppression; cytotoxicity; leukopenia; Crohn's disease;
KW autoimmune disease; rheumatoid arthritis; multiple sclerosis; cancer;
KW inflammatory disease; pancreatitis; inflammatory bowel disease.
XX
OS Mus musculus.
XX
PN WO200200721-A2.
XX
PD 03-JAN-2002.
XX
PF 26-JUN-2001; 2001WO-US20484.
XX
PR 26-JUN-2000; 2000US-214282P.
PR 29-JUN-2000; 2000US-214955P.
PR 08-FEB-2001; 2001US-267963P.
XX
PA (ZYMO) ZYMOGENETICS INC.
XX
PI Sprecher CA, Presnell SR, Gao Z, Whitmore TE, Kuijper JL;
PI Maurer MF;
XX
DR WPI; 2002-090519/12.
DR N-PSDB; ABA93843.
XX
PT Isolated polynucleotide encoding a cytokine receptor zcytor17 which is
PT useful for treating and diagnosing lymphoid, immune, inflammatory,
PT spleenic, blood or bone disorders -
XX
PS Example 21; Page 233-235; 235pp; English.
XX
CC The present invention describes a cytokine receptor designated zcytor17.
CC Zcytor17 has immunomodulatory, antiinflammatory, antiviral, cytostatic,
CC antirheumatic, antiarthritic and muscular activities. The zcytor17
CC proteins are useful for treating and diagnosing lymphoid, immune,
CC inflammatory, spleenic, blood or bone disorders. Agonists or
CC anti-zcytor17 antibodies are useful in stimulating cell-mediated
CC immunity and for stimulating lymphocyte proliferation, such as in the
CC treatment of infections involving immunosuppression, including certain
CC viral infections. They are also useful for inducting cytotoxicity and
CC for treating leukopenias. Antagonist of zcytor17 polypeptides are useful
CC for treating autoimmune diseases (e.g. rheumatoid arthritis and multiple
CC sclerosis), inflammatory diseases (e.g. Crohn's disease), cancer,
CC pancreatitis, and inflammatory bowel disease. Zcytor17 was mapped to
CC chromosome 5, specifically to the 5q11 chromosomal region. ABA93767 to
CC ABA93843 and ABB05730 to ABB05745 represent sequences used in the
CC exemplification of the present invention.
XX
SQ Sequence 547 AA;

Query Match 75.0%; Score 36; DB 23; Length 547;
Best Local Similarity 55.6%; Pred. No. 4.5e+02;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 KKLQILWKK 9
              :|:::||||
Db        258 RKVRLLWKK 266
```

Search completed: October 14, 2003, 14:41:32
Job time : 39.7917 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 14:33:23 ; Search time 12.75 Seconds
(without alignments)
29.866 Million cell updates/sec

Title: US-09-881-490-163
Perfect score: 48
Sequence: 1 KKLQILWKK 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	48	100.0	9	2	US-08-621-259A-163	Sequence 163, App
2	48	100.0	9	5	PCT-US95-09262-163	Sequence 163, App
3	43	89.6	8	2	US-08-621-259A-205	Sequence 205, App
4	43	89.6	9	2	US-08-621-259A-207	Sequence 207, App
5	35	72.9	303	4	US-09-205-258-818	Sequence 818, App
6	35	72.9	400	4	US-09-205-258-822	Sequence 822, App
7	35	72.9	453	4	US-09-252-991A-31268	Sequence 31268, A
8	34	70.8	18	4	US-09-450-315A-1	Sequence 1, Appli
9	34	70.8	20	4	US-09-000-692-1	Sequence 1, Appli
10	34	70.8	281	4	US-09-107-532A-6750	Sequence 6750, Ap
11	33	68.8	9	2	US-08-621-803-220	Sequence 220, App
12	33	68.8	9	2	US-08-621-259A-212	Sequence 212, App
13	33	68.8	9	3	US-09-217-352-220	Sequence 220, App
14	33	68.8	10	2	US-08-621-259A-198	Sequence 198, App
15	33	68.8	10	2	US-08-621-259A-199	Sequence 199, App
16	33	68.8	10	5	PCT-US95-09262-198	Sequence 198, App
17	33	68.8	10	5	PCT-US95-09262-199	Sequence 199, App
18	33	68.8	358	4	US-09-328-352-7579	Sequence 7579, Ap
19	33	68.8	822	4	US-09-198-452A-956	Sequence 956, App
20	32	66.7	33	2	US-08-653-632-63	Sequence 63, Appl
21	32	66.7	33	4	US-09-194-296-6	Sequence 6, Appli
22	32	66.7	644	4	US-09-107-532A-7277	Sequence 7277, Ap
23	32	66.7	1095	4	US-09-107-532A-3855	Sequence 3855, Ap
24	32	66.7	1464	3	US-09-045-360-2	Sequence 2, Appli
25	31	64.6	96	4	US-09-227-357-159	Sequence 159, App
26	31	64.6	185	4	US-09-529-157-6	Sequence 6, Appli
27	31	64.6	185	4	US-09-529-157-7	Sequence 7, Appli
28	31	64.6	221	4	US-09-495-406-23	Sequence 23, Appl
29	31	64.6	240	4	US-09-107-532A-4551	Sequence 4551, Ap
30	31	64.6	255	2	US-08-933-750C-7	Sequence 7, Appli
31	31	64.6	255	3	US-09-234-613-7	Sequence 7, Appli
32	31	64.6	388	1	US-08-429-742-4	Sequence 4, Appli
33	31	64.6	415	3	US-08-861-774E-26	Sequence 26, Appl
34	31	64.6	518	3	US-09-045-360-4	Sequence 4, Appli
35	31	64.6	686	4	US-09-653-839-8	Sequence 8, Appli
36	31	64.6	702	4	US-09-653-839-6	Sequence 6, Appli
37	31	64.6	723	4	US-09-653-839-4	Sequence 4, Appli
38	31	64.6	739	4	US-09-653-839-2	Sequence 2, Appli
39	31	64.6	1064	3	US-08-726-214-8	Sequence 8, Appli
40	31	64.6	1077	4	US-09-412-210-1	Sequence 1, Appli
41	31	64.6	1167	3	US-08-857-076-48	Sequence 48, Appl
42	30	62.5	14	1	US-08-311-611A-106	Sequence 106, App
43	30	62.5	14	1	US-08-372-783-106	Sequence 106, App
44	30	62.5	14	1	US-08-372-105-106	Sequence 106, App
45	30	62.5	14	1	US-08-306-473A-106	Sequence 106, App

ALIGNMENTS

```
RESULT 1
US-08-621-259A-163
; Sequence 163, Application US/08621259A
; Patent No. 5858974
;  GENERAL INFORMATION:
;    APPLICANT:  Little II, Roger G
;    APPLICANT:  Lim, Edward
;    APPLICANT:  Fadem, Mitchell B.
;    TITLE OF INVENTION:  Anti-Fungal Peptides
;    NUMBER OF SEQUENCES:  252
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McAndrews, Held & Malloy, Ltd.
;      STREET:  500 West Madison Street
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60661
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,259A
;      FILING DATE:  21-MAR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/504,841
;      FILING DATE:  20-JUL-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McNicholas, Janet M.
;      REGISTRATION NUMBER:  32,918
;      REFERENCE/DOCKET NUMBER:  11021US02
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/707-8889
;      TELEFAX:  312/707-9155
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  163:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.332"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1-9
;      OTHER INFORMATION:  /label= D-Amino Acids
;      OTHER INFORMATION:  /note= "Positions 1-9 are D-amino acids"
```

```
:
'KEY:  Modified-site
 ION:  C-Terminus
R INFORMATION:  /label= Amidation
R INFORMATION:  /note= "The C-Terminus is Amidated."
59A-163

.ch             100.0%;  Score 48;  DB 2;  Length 9;
il Similarity   100.0%;  Pred. No. 2.5e+05;
  9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

  1 KKLQILWKK 9
    |||||||||
  1 KKLQILWKK 9


09262-163
e 163, Application PC/TUS9509262
L INFORMATION:
ICANT:
E OF INVENTION:  Anti-Fungal Peptides
ER OF SEQUENCES:  206
ESPONDENCE ADDRESS:
DRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
REET:  6300 Sears Tower, 233 South Wacker Drive
ITY:  Chicago
TATE:  Illinois
OUNTRY:  United States of America
IP:  60606-6402
PUTER READABLE FORM:
EDIUM TYPE:  Floppy disk
OMPUTER:  IBM PC compatible
PERATING SYSTEM:  PC-DOS/MS-DOS
SOFTWARE:  PatentIn Release #1.0, Version #1.25
RRENT APPLICATION DATA:
APPLICATION NUMBER:  PCT/US95/09262
FILING DATE:
RIOR APPLICATION DATA:
APPLICATION NUMBER:  08/372,105
FILING DATE:  13-JAN-95
RIOR APPLICATION DATA:
APPLICATION NUMBER:  08/306,473
FILING DATE:  15-SEP-94
RIOR APPLICATION DATA:
APPLICATION NUMBER:  08/273,540
FILING DATE:  11-JUL-94
RIOR APPLICATION DATA:
APPLICATION NUMBER:  08/209,762
FILING DATE:  11-MAR-94
PRIOR APPLICATION DATA:
APPLICATION NUMBER:  08/183,222
FILING DATE:  14-JAN-94
PRIOR APPLICATION DATA:
APPLICATION NUMBER:  08/093,202
FILING DATE:  15-JUL-93
PRIOR APPLICATION DATA:
APPLICATION NUMBER:  08/03C,644
FILING DATE:  12-MAR-93
ATTORNEY/AGENT INFORMATION:
NAME:  Borun, Michael F.
REGISTRATION NUMBER:  25,447
REFERENCE/DOCKET NUMBER:  27129/10040
TELECOMMUNICATION INFORMATION:
TELEPHONE:  312/474-6300
TELEFAX:  312/474-0448
TELEX:  25-3856
INFORMATION FOR SEQ ID NO:  163:
SEQUENCE CHARACTERISTICS:
LENGTH:  9 amino acids
TYPE:  amino acid
TOPOLOGY:  linear
MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.332"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1-9
;      OTHER INFORMATION:  /label= D-Amino Acids
;      OTHER INFORMATION:  /note= "Positions 1-9 are D-amino acids"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated"
PCT-US95-09262-163

  Query Match             100.0%;  Score 48;  DB 5;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.5e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLQILWKK 9
              |||||||||
Db          1 KKLQILWKK 9


RESULT 3
US-08-621-259A-205
; Sequence 205, Application US/08621259A
; Patent No. 5858974
;  GENERAL INFORMATION:
;    APPLICANT:  Little II, Roger G
;    APPLICANT:  Lim, Edward
;    APPLICANT:  Fadem, Mitchell B.
;    TITLE OF INVENTION:  Anti-Fungal Peptides
;    NUMBER OF SEQUENCES:  252
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McAndrews, Held & Malloy, Ltd.
;      STREET:  500 West Madison Street
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60661
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,259A
;      FILING DATE:  21-MAR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/504,841
;      FILING DATE:  20-JUL-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McNicholas, Janet M.
;      REGISTRATION NUMBER:  32,918
;      REFERENCE/DOCKET NUMBER:  11021US02
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/707-8889
;      TELEFAX:  312/707-9155
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  205:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  8 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.374"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
```

```
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1-8
;      OTHER INFORMATION:  /label= D-Amino Acids
;      OTHER INFORMATION:  /note= "Positions 1-8 are D-Amino Acids."
US-08-621-259A-205

  Query Match             89.6%;  Score 43;  DB 2;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 2.5e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 KLQILWKK 9
              ||||||||
Db          1 KLQILWKK 8


RESULT 4
US-08-621-259A-207
; Sequence 207, Application US/08621259A
; Patent No. 5858974
;  GENERAL INFORMATION:
;    APPLICANT:  Little II, Roger G
;    APPLICANT:  Lim, Edward
;    APPLICANT:  Fadem, Mitchell B.
;    TITLE OF INVENTION:  Anti-Fungal Peptides
;    NUMBER OF SEQUENCES:  252
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McAndrews, Held & Malloy, Ltd.
;      STREET:  500 West Madison Street
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60661
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,259A
,      FILING DATE:  21-MAR-1996
,    PRIOR APPLICATION DATA:
,      APPLICATION NUMBER:  08/504,841
,      FILING DATE:  20-JUL-1995
,    ATTORNEY/AGENT INFORMATION:
,      NAME:  McNicholas, Janet M.
,      REGISTRATION NUMBER:  32,918
,      REFERENCE/DOCKET NUMBER:  11021US02
,    TELECOMMUNICATION INFORMATION:
,      TELEPHONE:  312/707-8889
       TELEFAX:  312/707-9155
       TELEX:
   INFORMATION FOR SEQ ID NO:  207:
     SEQUENCE CHARACTERISTICS:
       LENGTH:  9 amino acids
       TYPE:  amino acid
       TOPOLOGY:  linear
     MOLECULE TYPE:  peptide
     FEATURE:
       NAME/KEY:  misc_feature
       OTHER INFORMATION:  "XMP.376"
     FEATURE:
       NAME/KEY:  Modified-site
       LOCATION:  C-Terminus
       OTHER INFORMATION:  /label= Amidation
       OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
     FEATURE:
       NAME/KEY:  Modified-site
       LOCATION:  1-9
       OTHER INFORMATION:  /label= D-Amino Acids
       OTHER INFORMATION:  /note= "Positions 1-9 are D-Amino Acids."
 S-08-621-259A-207

  Query Match             89.6%;  Score 43;  DB 2;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 2.5e+05;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKLQILWKK 9
              ||||| |||
Db          1 KKLQIAWKK 9


RESULT 5
US-09-205-258-818
; Sequence 818, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
;  APPLICANT: Young et al.
;  TITLE OF INVENTION: 207 Human Secreted Proteins
;  FILE REFERENCE: PZ007P1
;  CURRENT APPLICATION NUMBER: US/09/205,258
;  CURRENT FILING DATE: 1998-12-04
;  EARLIER APPLICATION NUMBER: PCT/US98/11422
;  EARLIER FILING DATE: 1998-06-04
;  EARLIER APPLICATION NUMBER: 60/048,885
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,375
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,881
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,880
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,896
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,020
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,876
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,895
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,884
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,894
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,971
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,964
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,882
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,899
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,893
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,900
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,901
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,892
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,915
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,019
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,970
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,972
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,916
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,373
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,875
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,374
```

; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 818
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-205-258-818

Query Match 72.9%; Score 35; DB 4; Length 303;
Best Local Similarity 75.0%; Pred. No. 1.3e-02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 KKLQILWK 8
              ::|||||||
Db        199 QQLQILWK 206
```

RESULT 6
US-09-205-258-822
; Sequence 822, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: PZ007P1
; CURRENT APPLICATION NUMBER: US/09/205,258
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 822
; LENGTH: 400
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-205-258-822

Query Match 72.9%; Score 35; DB 4; Length 400;
Best Local Similarity 75.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 KKLQILWK 8
              ::|||||
Db        199 QQLQILWK 206


RESULT 7
US-09-252-991A-31268
; Sequence 31268, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 31268
;   LENGTH: 453
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31268

  Query Match             72.9%;  Score 35;  DB 4;  Length 453;
  Best Local Similarity   66.7%;  Pred. No. 1.9e+02;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKLQILWKK 9
              |||. :|:|
Db        158 KKLQEIWEK 166


RESULT 8
US-09-450-315A-1
; Sequence 1, Application US/09450315A
; Patent No. 6379966
; GENERAL INFORMATION:
;  APPLICANT: WOLFF, JON A
;  APPLICANT:  Monahan, Sean D
;  APPLICANT:  Slattum, Paul M
;  APPLICANT:  Hagstrom, James E
;  APPLICANT:  Budker, Vladimir G
;  APPLICANT:  Rozema, David B
;  TITLE OF INVENTION: INTRAVASCULAR DELIVERY OF NON-VIRAL NUCLEIC ACID
;  FILE REFERENCE: Mirus.013.03
;  CURRENT APPLICATION NUMBER: US/09/450,315A
;  CURRENT FILING DATE:  1999-11-29
;  PRIOR APPLICATION NUMBER: 60/121730
;  PRIOR FILING DATE: 1999-02-26
;  PRIOR APPLICATION NUMBER: 60/146564
;  PRIOR FILING DATE: 1999-07-30
;  NUMBER OF SEQ ID NOS: 1
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 18
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: synthesized
;   OTHER INFORMATION: amphipathic peptide with alternating hydrophilic
;   OTHER INFORMATION: and hydrophobic residues.
US-09-450-315A-1

  Query Match             70.8%;  Score 34;  DB 4;  Length 18;
  Best Local Similarity   77.8%;  Pred. No. 15;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 KKLQILWKK 9
              |||  ||||
Db          4 KKLLKLWKK 12


RESULT 9
US-09-000-692-1
; Sequence 1, Application US/09000692
; Patent No. 6339067
; GENERAL INFORMATION:
;  APPLICANT: WOLFF, JON A
;  APPLICANT: HAGSTROM, JAMES E
;  APPLICANT: BUDKER, VLADIMIR G
;  APPLICANT: TRUBETSKOY, VLADIMIR S
;  APPLICANT: SLATTUM, PAUL M
;  APPLICANT: HANSON, LISA J
;  TITLE OF INVENTION: A PROCESS OF MAKING A COMPOUND BY FORMING A POLYMER
;  TITLE OF INVENTION: FROM A TEMPLATE  DRUG
;  FILE REFERENCE: TPCIP000692
;  CURRENT APPLICATION NUMBER: US/09/000,692
;  CURRENT FILING DATE: 1997-12-30
;  EARLIER APPLICATION NUMBER: 08/778657
;  EARLIER FILING DATE: 1997-01-03
;  NUMBER OF SEQ ID NOS: 1
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 20
;   TYPE: PRT
;   ORGANISM: Unknown
;   FEATURE:
;   OTHER INFORMATION: Description of Unknown Organism:AMPHIPATHIC
US-09-000-692-1

  Query Match             70.8%;  Score 34;  DB 4;  Length 20;
  Best Local Similarity   77.8%;  Pred. No. 17;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 KKLQILWKK 9
              |||  ||||
Db          5 KKLLKLWKK 13


RESULT 10
US-09-107-532A-6750
; Sequence 6750, Application US/09107532A
; Patent No. 6583275
;   GENERAL INFORMATION:
;        APPLICANT: Lynn A Doucette-Stamm  and David Bush
;        TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
;                            ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
;        NUMBER OF SEQUENCES: 7310
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: GENOME THERAPEUTICS CORPORATION
;             STREET: 100 Beaver Street
;             CITY: Waltham
;             STATE: Massachusetts
;             COUNTRY: USA
;             ZIP: 02354
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: CD/ROM ISO9660
;             COMPUTER: PC
;             OPERATING SYSTEM: <Unknown>
;             SOFTWARE: ASCII
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/107,532A
;             FILING DATE: 30-Jun-1998
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 60/085,598
;             FILING DATE: 14 May 1998
;             APPLICATION NUMBER: 60/051571
;             FILING DATE: July 2, 1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Ariniello, Pamela Deneke
;             REGISTRATION NUMBER: 40,489
;             REFERENCE/DOCKET NUMBER: GTC-012
```

```
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (781)893-5007
;            TELEFAX: (781)893-8277
;   INFORMATION FOR SEQ ID NO: 6750:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 281 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       HYPOTHETICAL: YES
;       ORIGINAL SOURCE:
;            ORGANISM: Enterococcus faecium
;       FEATURE:
;            NAME/KEY:  misc_feature
;            LOCATION: (B) LOCATION 1...281
;       SEQUENCE DESCRIPTION: SEQ ID NO: 6750:
US-09-107-532A-6750

  Query Match             70.8%;  Score 34;  DB 4;  Length 281;
  Best Local Similarity   55.6%;  Pred. No. 1.8e+02;
  Matches    5;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKLQILWKK 9
              :|:|:|| |
Db         27 QKVQVLWPK 35


RESULT 11
US-08-621-803-220
; Sequence 220, Application US/08621803
; Patent No. 5851802
;  GENERAL INFORMATION:
;    APPLICANT:  Better, Marc D.
;    TITLE OF INVENTION:  Methods for Recombinant Microbial Production of
;    TITLE OF INVENTION:  Fusion Proteins and BPI-Derived Peptides
;    NUMBER OF SEQUENCES:  265
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,803
;      FILING DATE:  22-MAR-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25,447
;      REFERENCE/DOCKET NUMBER:  27129/33199
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  220:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.381"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-08-621-803-220

  Query Match             68.8%;  Score 33;  DB 2;  Length 9;
  Best Local Similarity   77.8%;  Pred. No. 2.5e+05;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKLQILWKK 9
              ||||:| ||
Db          1 KKLQLLLKK 9


RESULT 12
US-08-621-259A-212
; Sequence 212, Application US/08621259A
; Patent No. 5858974
;  GENERAL INFORMATION:
;    APPLICANT:  Little II, Roger G
;    APPLICANT:  Lim, Edward
;    APPLICANT:  Fadem, Mitchell B.
;    TITLE OF INVENTION:  Anti-Fungal Peptides
;    NUMBER OF SEQUENCES:  252
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McAndrews, Held & Malloy, Ltd.
;      STREET:  500 West Madison Street
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60661
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,259A
;      FILING DATE:  21-MAR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/504,841
;      FILING DATE:  20-JUL-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McNicholas, Janet M.
;      REGISTRATION NUMBER:  32,918
;      REFERENCE/DOCKET NUMBER:  11021US02
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/707-8889
;      TELEFAX:  312/707-9155
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  212:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.381"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-08-621-259A-212

  Query Match             68.8%;  Score 33;  DB 2;  Length 9;
  Best Local Similarity   77.8%;  Pred. No. 2.5e+05;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKLQILWKK 9
              ||||:| ||
Db          1 KKLQLLLKK 9


RESULT 13
```

```
US-09-217-352-220
; Sequence 220, Application US/09217352
; Patent No. 6274344
;  GENERAL INFORMATION:
;    APPLICANT:  Better, Marc D.
;    TITLE OF INVENTION:  Methods for Recombinant Microbial Production of
;    TITLE OF INVENTION:  Fusion Proteins and BPI-Derived Peptides
;    NUMBER OF SEQUENCES:  265
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/217,352
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/621,803
;      FILING DATE:  22-MAR-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25,447
;      REFERENCE/DOCKET NUMBER:  27129/33199
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  220:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.381"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-09-217-352-220

  Query Match             68.8%;  Score 33;  DB 3;  Length 9;
  Best Local Similarity   77.8%;  Pred. No. 2.5e+05;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKLQILWKK 9
              |||||:| ||
Db          1 KKLQLLLKK 9


RESULT 14
US-08-621-259A-198
; Sequence 198, Application US/08621259A
; Patent No. 5858974
;  GENERAL INFORMATION:
;    APPLICANT:  Little II, Roger G
;    APPLICANT:  Lim, Edward
;    APPLICANT:  Fadem, Mitchell B.
;    TITLE OF INVENTION:  Anti-Fungal Peptides
;    NUMBER OF SEQUENCES:  252
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McAndrews, Held & Malloy, Ltd.
;      STREET:  500 West Madison Street
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60661
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,259A
;      FILING DATE:  21-MAR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/504,841
;      FILING DATE:  20-JUL-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McNicholas, Janet M.
;      REGISTRATION NUMBER:  32,918
;      REFERENCE/DOCKET NUMBER:  11021US02
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/707-8889
;      TELEFAX:  312/707-9155
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  198:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.367"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1-10
;      OTHER INFORMATION:  /label= D-Amino Acids
;      OTHER INFORMATION:  /note= "Positions 1-10 are D-amino acids"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-08-621-259A-198

  Query Match             68.8%;  Score 33;  DB 2;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LQILWK 8
              ||||||
Db          5 LQILWK 10


RESULT 15
US-08-621-259A-199
; Sequence 199, Application US/08621259A
; Patent No. 5858974
;  GENERAL INFORMATION:
;    APPLICANT:  Little II, Roger G
;    APPLICANT:  Lim, Edward
;    APPLICANT:  Fadem, Mitchell B.
;    TITLE OF INVENTION:  Anti-Fungal Peptides
;    NUMBER OF SEQUENCES:  252
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McAndrews, Held & Malloy, Ltd.
;      STREET:  500 West Madison Street
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60661
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
```

```
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,259A
;      FILING DATE:  21-MAR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/504,841
;      FILING DATE:  20-JUL-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McNicholas, Janet M.
;      REGISTRATION NUMBER:  32,918
;      REFERENCE/DOCKET NUMBER:  11021US02
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/707-8889
;      TELEFAX:  312/707-9155
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  199:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.368"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1
;      OTHER INFORMATION:  /label= Acetylated
;      OTHER INFORMATION:  /note= "Position 1 is acetylated."
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1-10
;      OTHER INFORMATION:  /label= D-Amino Acids
;      OTHER INFORMATION:  /note= "Positions 1-10 are D-amino acids"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-08-621-259A-199

  Query Match             68.8%;  Score 33;  DB 2;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LQILWK 8
              ||||||
Db          5 LQILWK 10


Search completed: October 14, 2003, 14:42:46
Job time : 13.75 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 14:34:14 ; Search time 21.9375 Seconds
(without alignments)
66.104 Million cell updates/sec

Title: US-09-881-490-163
Perfect score: 48
Sequence: 1 KKLQIILWKK 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 600653 seqs, 161128416 residues

Total number of hits satisfying chosen parameters: 600653

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	48	100.0	9	9	US-09-881-490-163	Sequence 163, App
2	38	79.2	325	15	US-10-082-830-193	Sequence 193, App
3	36	75.0	547	11	US-09-892-949-93	Sequence 93, Appl
4	36	75.0	662	11	US-09-892-949-57	Sequence 57, Appl
5	36	75.0	726	11	US-09-972-708-12	Sequence 12, Appl
6	35	72.9	231	12	US-10-107-857-48	Sequence 48, Appl
7	35	72.9	303	12	US-09-933-767-818	Sequence 818, App
8	35	72.9	303	15	US-10-023-282-818	Sequence 818, App
9	35	72.9	400	12	US-09-933-767-822	Sequence 822, App
10	35	72.9	400	15	US-10-023-282-822	Sequence 822, App
11	35	72.9	738	12	US-10-175-270-2	Sequence 2, Appli
12	35	72.9	738	12	US-10-107-857-30	Sequence 30, Appl
13	35	72.9	740	12	US-10-325-430-15	Sequence 15, Appl
14	35	72.9	795	12	US-10-107-857-21	Sequence 21, Appl
15	35	72.9	855	12	US-10-107-857-45	Sequence 45, Appl
16	34	70.8	18	9	US-09-447-966-1	Sequence 1, Appli
17	34	70.8	18	12	US-10-085-378-2	Sequence 2, Appli
18	34	70.8	49	9	US-09-864-761-41571	Sequence 41571, A
19	34	70.8	560	10	US-09-978-295A-420	Sequence 420, App
20	34	70.8	560	10	US-09-978-697-420	Sequence 420, App
21	34	70.8	560	10	US-09-978-192A-420	Sequence 420, App
22	34	70.8	560	10	US-09-999-832A-420	Sequence 420, App
23	34	70.8	560	11	US-09-978-189-420	Sequence 420, App
24	34	70.8	560	11	US-09-978-608A-420	Sequence 420, App
25	34	70.8	560	11	US-09-978-585A-420	Sequence 420, App
26	34	70.8	560	11	US-09-978-191A-420	Sequence 420, App
27	34	70.8	560	11	US-09-978-403A-420	Sequence 420, App
28	34	70.8	560	11	US-09-978-564A-420	Sequence 420, App
29	34	70.8	560	11	US-09-999-833A-420	Sequence 420, App
30	34	70.8	560	11	US-09-981-915A-420	Sequence 420, App
31	34	70.8	560	11	US-09-978-824-420	Sequence 420, App
32	34	70.8	560	11	US-09-918-585A-420	Sequence 420, App
33	34	70.8	560	11	US-09-978-423A-420	Sequence 420, App
34	34	70.8	560	11	US-09-978-193A-420	Sequence 420, App
35	34	70.8	560	11	US-09-999-830A-420	Sequence 420, App
36	34	70.8	560	11	US-09-978-757A-420	Sequence 420, App
37	34	70.8	560	11	US-09-978-187B-420	Sequence 420, App
38	34	70.8	560	11	US-09-978-643A-420	Sequence 420, App
39	34	70.8	560	12	US-09-978-375A-420	Sequence 420, App
40	34	70.8	560	12	US-09-978-188A-420	Sequence 420, App
41	34	70.8	560	12	US-09-978-298A-420	Sequence 420, App
42	34	70.8	560	12	US-10-143-031A-420	Sequence 420, App
43	34	70.8	560	12	US-10-002-967A-420	Sequence 420, App
44	34	70.8	560	12	US-10-017-083A-420	Sequence 420, App
45	34	70.8	560	12	US-10-143-030A-420	Sequence 420, App

ALIGNMENTS

```
RESULT 1
US-09-881-490-163
; Sequence 163, Application US/09881490
; Patent No. US20020077298A1
;  GENERAL INFORMATION:
;       APPLICANT: Little II, Roger G.
;                  Lim, Edward
;                  Fadem, Mitchell B.
;       TITLE OF INVENTION: Anti-Fungal Peptides
;       NUMBER OF SEQUENCES: 211
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: McAndrews, Held & Malloy, Ltd.
;            STREET: 500 West Madison Street, 34th FloorDrive
;            CITY: Chicago
;            STATE: Illinois
;            COUNTRY: United States of America
;            ZIP: 60661
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/881,490
;            FILING DATE: 14-Jun-2001
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/119,858
;            FILING DATE: <Unknown>
;            APPLICATION NUMBER: 08/372,105
;            FILING DATE: 13-JAN-95
;            APPLICATION NUMBER: 08/306,473
;            FILING DATE: 15-SEP-94
;            APPLICATION NUMBER: 08/273,540
;            FILING DATE: 11-JUL-94
;            APPLICATION NUMBER: 08/209,762
;            FILING DATE: 11-MAR-94
;            APPLICATION NUMBER: 08/183,222
;            FILING DATE: 14-JAN-94
```

```
;             APPLICATION NUMBER: 08/093,202
;             FILING DATE: 15-JUL-93
;             APPLICATION NUMBER: 08/030,644
;             FILING DATE: 12-MAR-93
;        ATTORNEY/AGENT INFORMATION:
;             NAME: McNicholas, Janet M.
;             REGISTRATION NUMBER: 32,918
;             REFERENCE/DOCKET NUMBER: 100-238/11021US01
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 312/707-8889
;             TELEFAX: 312/707-9155
;             TELEX: 650 388-1248
;   INFORMATION FOR SEQ ID NO: 163:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 9 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        FEATURE:
;             NAME/KEY:  misc_feature
;             OTHER INFORMATION:  "XMP.332"
;        FEATURE:
;             NAME/KEY:  Modified-site
;             LOCATION:  1-9
;             OTHER INFORMATION:  /label= D-Amino Acids
;             /note= "Positions 1-9 are D-amino acids"
;        FEATURE:
;             NAME/KEY:  Modified-site
;             LOCATION:  C-Terminus
;             OTHER INFORMATION:  /label= Amidation
;             /note= "The C-Terminus is Amidated"
;        SEQUENCE DESCRIPTION: SEQ ID NO: 163:
US-09-881-490-163

  Query Match             100.0%;  Score 48;  DB 9;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 5.3e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLQILWKK 9
              |||||||||
Db          1 KKLQILWKK 9


RESULT 2
US-10-082-830-193
; Sequence 193, Application US/10082830
; Publication No. US20030077604A1
; GENERAL INFORMATION:
;  APPLICANT: Sun, Yongming
;  APPLICANT:  Recipon, Herve
;  APPLICANT:  Salceda, Susana
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Turner, Leah
;  TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific
;  TITLE OF INVENTION:  Genes and Proteins
;  FILE REFERENCE: DEX-0249
;  CURRENT APPLICATION NUMBER: US/10/082,830
;  CURRENT FILING DATE:  2001-10-23
;  PRIOR APPLICATION NUMBER: 60/243,802
;  PRIOR FILING DATE: 2000-10-27
;  NUMBER OF SEQ ID NOS: 282
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 193
;   LENGTH: 325
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-082-830-193

  Query Match             79.2%;  Score 38;  DB 15;  Length 325;
  Best Local Similarity   75.0%;  Pred. No. 98;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 KLQILWKK 9
              |||::|||
Db         78 KLQLVWKK 85


RESULT 3
US-09-892-949-93
; Sequence 93, Application US/09892949
; Publication No. US20030096339A1
; GENERAL INFORMATION:
;  APPLICANT: Sprecher, Cindy A.
;  APPLICANT:  Presnell, Scott R.
;  APPLICANT:  Gao, Zeren
;  APPLICANT:  Whitmore, Theodore E.
;  APPLICANT:  Kuijper, Joseph L.
;  APPLICANT:  Maurer, Mark F.
;  TITLE OF INVENTION: CYTOKINE RECEPTOR ZCYTOR17
;  FILE REFERENCE: 00-42
;  CURRENT APPLICATION NUMBER: US/09/892,949
;  CURRENT FILING DATE:  2001-06-26
;  PRIOR APPLICATION NUMBER: US 60/214,282
;  PRIOR FILING DATE: 2000-06-26
;  PRIOR APPLICATION NUMBER: US 60/214,955
;  PRIOR FILING DATE: 2000-06-29
;  PRIOR APPLICATION NUMBER: US 60/267,963
;  PRIOR FILING DATE: 2001-08-02
;  NUMBER OF SEQ ID NOS: 93
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 93
;   LENGTH: 547
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-892-949-93

  Query Match             75.0%;  Score 36;  DB 11;  Length 547;
  Best Local Similarity   55.6%;  Pred. No. 3.2e+02;
  Matches    5;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLQILWKK 9
              :|:::||||
Db        258 RKVRLLWKK 266


RESULT 4
US-09-892-949-57
; Sequence 57, Application US/09892949
; Publication No. US20030096339A1
; GENERAL INFORMATION:
;  APPLICANT: Sprecher, Cindy A.
;  APPLICANT:  Presnell, Scott R.
;  APPLICANT:  Gao, Zeren
;  APPLICANT:  Whitmore, Theodore E.
;  APPLICANT:  Kuijper, Joseph L.
;  APPLICANT:  Maurer, Mark F.
;  TITLE OF INVENTION: CYTOKINE RECEPTOR ZCYTOR17
;  FILE REFERENCE: 00-42
;  CURRENT APPLICATION NUMBER: US/09/892,949
;  CURRENT FILING DATE:  2001-06-26
;  PRIOR APPLICATION NUMBER: US 60/214,282
;  PRIOR FILING DATE: 2000-06-26
;  PRIOR APPLICATION NUMBER: US 60/214,955
;  PRIOR FILING DATE: 2000-06-29
;  PRIOR APPLICATION NUMBER: US 60/267,963
;  PRIOR FILING DATE: 2001-08-02
;  NUMBER OF SEQ ID NOS: 93
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 57
;   LENGTH: 662
;   TYPE: PRT
;   ORGANISM: mus musculus
US-09-892-949-57

  Query Match             75.0%;  Score 36;  DB 11;  Length 662;
  Best Local Similarity   55.6%;  Pred. No. 3.8e+02;
```

```
  Matches    5;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLQILWKK 9
              :|:::||||
Db        258 RKVRLLWKK 266


RESULT 5
US-09-972-708-12
; Sequence 12, Application US/09972708
; Publication No. US20030059871A1
; GENERAL INFORMATION:
;  APPLICANT: Immunex Corporation
;  APPLICANT:  Cosman, David J.
;  APPLICANT:  Mosley, Bruce A.
;  APPLICANT:  Bird, Timothy A.
;  APPLICANT:  DuBose, Robert F.
;  APPLICANT:  Wiley, Steven R.
;  TITLE OF INVENTION: HEMATOPOIETIN RECEPTORS HPR1 AND HPR2
;  FILE REFERENCE: 3160-B
;  CURRENT APPLICATION NUMBER: US/09/972,708
;  CURRENT FILING DATE:  2001-10-05
;  NUMBER OF SEQ ID NOS: 29
;  SOFTWARE: Patentin version 3.1
; SEQ ID NO 12
;   LENGTH: 726
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-972-708-12

  Query Match             75.0%;  Score 36;  DB 11;  Length 726;
  Best Local Similarity   55.6%;  Pred. No. 4.1e-02;
  Matches    5;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLQILWKK 9
              :|:::||||
Db        241 RKVRLLWKK 249


RESULT 6
US-10-107-857-48
; Sequence 48, Application US/10107857
; Publication No. US20030170784A1
; GENERAL INFORMATION:
;  APPLICANT: Fraser, Christopher C.
;  TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
;  FILE REFERENCE: 07334-251001
;  CURRENT APPLICATION NUMBER: US/10/107,857
;  CURRENT FILING DATE:  2002-03-26
;  PRIOR APPLICATION NUMBER: 09/514,009
;  PRIOR FILING DATE: 2000-02-25
;  NUMBER OF SEQ ID NOS: 51
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 48
;   LENGTH: 231
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-107-857-48

  Query Match             72.9%;  Score 35;  DB 12;  Length 231;
  Best Local Similarity   75.0%;  Pred. No. 2.2e+02;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLQILWK 8
              ::||||||
Db        127 QQLQILWK 134


RESULT 7
US-09-933-767-818
; Sequence 818, Application US/09933767
; Publication No. US20030181692A1
; GENERAL INFORMATION:
;  APPLICANT: Ni et al.
;  TITLE OF INVENTION: 207 Human Secreted Proteins
;  FILE REFERENCE: PZ007P2
;  CURRENT APPLICATION NUMBER: US/09/933,767
;  CURRENT FILING DATE:  2001-08-22
;  PRIOR APPLICATION NUMBER: PCT/US01/05614
;  PRIOR FILING DATE: 2001-02-21
;  PRIOR APPLICATION NUMBER: 60/184,836
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: 60/193,170
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: 09/205,258
;  PRIOR FILING DATE: 1998-12-04
;  PRIOR APPLICATION NUMBER: PCT/US98/11422
;  PRIOR FILING DATE: 1998-06-04
;  PRIOR APPLICATION NUMBER: 60/048,885
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/049,375
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,881
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,880
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,896
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/049,020
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,876
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,895
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,884
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,894
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,971
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,964
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,882
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,899
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,893
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,900
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,901
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,892
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,915
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/049,019
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,970
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,972
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,916
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/049,373
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,875
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/049,374
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,917
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,949
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,974
;  PRIOR FILING DATE: 1997-06-06
```

; PRIOR APPLICATION NUMBER: 60/048,883
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,897
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,898
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,962
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,963
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,877
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,878
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/068,054
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,064
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,053
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/070,923
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/073,160
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/073,159
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/073,165
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/073,164
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/085,925
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,921
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,923
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,922
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/092,921
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/094,657
; PRIOR FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1245
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 818
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-933-767-818

Query Match 72.9%; Score 35; DB 12; Length 303;
Best Local Similarity 75.0%; Pred. No. 2.8e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 KKLQILWK 8
              ::|||||||
Db        199 QQLQILWK 206
```

RESULT 8
US-10-023-282-818
; Sequence 818, Application US/10023282
; Publication No. US20030092893A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: PZ007P1
; CURRENT APPLICATION NUMBER: US/10/023,282
; CURRENT FILING DATE: 2001-12-20
; EARLIER APPLICATION NUMBER: 09/205,258
; EARLIER FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04

; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923

```
;  EARLIER FILING DATE: 1997-12-18
;  EARLIER APPLICATION NUMBER: 60/092,921
;  EARLIER FILING DATE: 1998-07-15
;  EARLIER APPLICATION NUMBER: 60/094,657
;  EARLIER FILING DATE: 1998-07-30
;  NUMBER OF SEQ ID NOS: 1227
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 818
;   LENGTH: 303
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-023-282-818

  Query Match             72.9%;  Score 35;  DB 15;  Length 303;
  Best Local Similarity   75.0%;  Pred. No. 2.8e-02;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLQILWK 8
              ::||||||
Db        199 QQLQILWK 206


RESULT 9
US-09-933-767-822
; Sequence 822, Application US/09933767
; Publication No. US20030181692A1
; GENERAL INFORMATION:
;  APPLICANT: Ni et al.
;  TITLE OF INVENTION: 207 Human Secreted Proteins
;  FILE REFERENCE: PZ007P2
;  CURRENT APPLICATION NUMBER: US/09/933,767
;  CURRENT FILING DATE:  2001-08-22
;  PRIOR APPLICATION NUMBER: PCT/US01/05614
;  PRIOR FILING DATE: 2001-02-21
;  PRIOR APPLICATION NUMBER: 60/184,836
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: 60/193,170
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: 09/205,258
;  PRIOR FILING DATE: 1998-12-04
;  PRIOR APPLICATION NUMBER: PCT/US98/11422
;  PRIOR FILING DATE: 1998-06-04
;  PRIOR APPLICATION NUMBER: 60/048,885
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/049,375
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,881
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,880
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,896
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/049,020
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,876
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,895
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,884
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,894
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,971
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,964
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,882
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,899
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,893
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,900
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,901
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,892
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,915
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/049,019
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,970
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,972
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,916
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/049,373
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,875
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/049,374
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,917
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,949
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,974
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,883
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,897
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,898
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,962
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,963
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,877
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,878
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/068,054
;  PRIOR FILING DATE: 1997-12-18
;  PRIOR APPLICATION NUMBER: 60/068,064
;  PRIOR FILING DATE: 1997-12-18
;  PRIOR APPLICATION NUMBER: 60/068,053
;  PRIOR FILING DATE: 1997-12-18
;  PRIOR APPLICATION NUMBER: 60/070,923
;  PRIOR FILING DATE: 1997-12-18
;  PRIOR APPLICATION NUMBER: 60/073,160
;  PRIOR FILING DATE: 1998-01-30
;  PRIOR APPLICATION NUMBER: 60/073,159
;  PRIOR FILING DATE: 1998-01-30
;  PRIOR APPLICATION NUMBER: 60/073,165
;  PRIOR FILING DATE: 1998-01-30
;  PRIOR APPLICATION NUMBER: 60/073,164
;  PRIOR FILING DATE: 1998-01-30
;  PRIOR APPLICATION NUMBER: 60/085,925
;  PRIOR FILING DATE: 1998-05-18
;  PRIOR APPLICATION NUMBER: 60/085,921
;  PRIOR FILING DATE: 1998-05-18
;  PRIOR APPLICATION NUMBER: 60/085,923
;  PRIOR FILING DATE: 1998-05-18
;  PRIOR APPLICATION NUMBER: 60/085,922
;  PRIOR FILING DATE: 1998-05-18
;  PRIOR APPLICATION NUMBER: 60/092,921
;  PRIOR FILING DATE: 1998-07-15
;  PRIOR APPLICATION NUMBER: 60/094,657
;  PRIOR FILING DATE: 1998-07-30
;  NUMBER OF SEQ ID NOS: 1245
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 822
;   LENGTH: 400
```

```
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-933-767-822

  Query Match             72.9%;  Score 35;  DB 12;  Length 400;
  Best Local Similarity   75.0%;  Pred. No. 3.6e+02;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLQILWK 8
              ::|||||||
Db        199 QQLQILWK 206


RESULT 10
US-10-023-282-822
; Sequence 822, Application US/10023282
; Publication No. US20030092893A1
; GENERAL INFORMATION:
;  APPLICANT: Young et al.
;  TITLE OF INVENTION: 207 Human Secreted Proteins
;  FILE REFERENCE: PZ007P1
;  CURRENT APPLICATION NUMBER: US/10/023,282
;  CURRENT FILING DATE: 2001-12-20
;  EARLIER APPLICATION NUMBER: 09/205,258
;  EARLIER FILING DATE: 1998-12-04
;  EARLIER APPLICATION NUMBER: PCT/US98/11422
;  EARLIER FILING DATE: 1998-06-04
;  EARLIER APPLICATION NUMBER: 60/048,885
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,375
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,881
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,880
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,896
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,020
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,876
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,895
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,884
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,894
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,971
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,964
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,882
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,899
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,893
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,900
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,901
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,892
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,915
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,019
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,970
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,972
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,916
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,373
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,875
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,374
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,917
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,949
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,974
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,883
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,897
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,898
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,962
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,963
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,877
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,878
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/070,923
;  EARLIER FILING DATE: 1997-12-18
;  EARLIER APPLICATION NUMBER: 60/092,921
;  EARLIER FILING DATE: 1998-07-15
;  EARLIER APPLICATION NUMBER: 60/094,657
;  EARLIER FILING DATE: 1998-07-30
;  NUMBER OF SEQ ID NOS: 1227
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 822
;   LENGTH: 400
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-023-282-822

  Query Match             72.9%;  Score 35;  DB 15;  Length 400;
  Best Local Similarity   75.0%;  Pred. No. 3.6e-02;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLQILWK 8
              ::|||||||
Db        199 QQLQILWK 206


RESULT 11
US-10-175-270-2
; Sequence 2, Application US/10175270
; Publication No. US20030166548A1
; GENERAL INFORMATION:
;  APPLICANT: PETERSON, JOHNNY W.
;  APPLICANT:  SAINI, SHAMSHER S.
;  APPLICANT:  WOOD, THOMAS G.
;  APPLICANT:  CHOPRA, ASHOK K.
;  TITLE OF INVENTION: ANTI-INFLAMMATORY AND OTHER THERAPEUTIC PROPHYLACTIC OR DIAGNC
;  TITLE OF INVENTION:  SYNTHETIC MELITTIN AND NE
;  TITLE OF INVENTION:  W RELATED PEPTIDES
;  FILE REFERENCE: UTSG:213
;  CURRENT APPLICATION NUMBER: US/10/175,270
;  CURRENT FILING DATE:  2002-06-19
;  PRIOR APPLICATION NUMBER: US/09/458,397
;  PRIOR FILING DATE: 1999-12-10
;  PRIOR APPLICATION NUMBER: PCT/US98/12340
;  PRIOR FILING DATE: 1998-06-11
;  NUMBER OF SEQ ID NOS: 13
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 738
;   TYPE: PRT
```

```
;   ORGANISM: Human
US-10-175-270-2

  Query Match             72.9%;  Score 35;  DB 12;  Length 738;
  Best Local Similarity   75.0%;  Pred. No. 6.1e+02;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLQILWK 8
              ::||||||
Db        537 QQLQILWK 544


RESULT 12
US-10-107-857-30
; Sequence 30, Application US/10107857
; Publication No. US20030170784A1
; GENERAL INFORMATION:
;  APPLICANT: Fraser, Christopher C.
;  TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
;  FILE REFERENCE: 07334-251001
;  CURRENT APPLICATION NUMBER: US/10/107,857
;  CURRENT FILING DATE:  2002-03-26
;  PRIOR APPLICATION NUMBER: 09/514,009
;  PRIOR FILING DATE: 2000-02-25
;  NUMBER OF SEQ ID NOS: 51
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 30
;   LENGTH: 738
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-107-857-30

  Query Match             72.9%;  Score 35;  DB 12;  Length 738;
  Best Local Similarity   75.0%;  Pred. No. 6.1e+02;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLQILWK 8
              ::||||||
Db        537 QQLQILWK 544


RESULT 13
US-10-325-430-15
; Sequence 15, Application US/10325430
; Publication No. US20030153525A1
; GENERAL INFORMATION:
;  APPLICANT: Millennium Pharmaceuticals, Inc
;  APPLICANT:  Silos-Santiago, Inmaculada
;  APPLICANT:  Rosenfeld, Julie Beth
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS IN TREATING
;  TITLE OF INVENTION:  PAIN AND PAINFUL DISORDERS USING 1465, 1587, 2146, 2207,
;  TITLE OF INVENTION:  32838,336 and 52908
;  FILE REFERENCE: MP101-294PIRNM
;  CURRENT APPLICATION NUMBER: US/10/325,430
;  CURRENT FILING DATE:  2002-12-19
;  PRIOR APPLICATION NUMBER: US 60/341,953
;  PRIOR FILING DATE: 2001-12-19
;  NUMBER OF SEQ ID NOS: 21
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15
;   LENGTH: 740
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-10-325-430-15

  Query Match             72.9%;  Score 35;  DB 12;  Length 740;
  Best Local Similarity   66.7%;  Pred. No. 6.1e+02;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 KKLQILWKK 9
              ||:|  |||
Db         84 KKIQTQWKK 92


RESULT 14
US-10-107-857-21
; Sequence 21, Application US/10107857
; Publication No. US20030170784A1
; GENERAL INFORMATION:
;  APPLICANT: Fraser, Christopher C.
;  TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
;  FILE REFERENCE: 07334-251001
;  CURRENT APPLICATION NUMBER: US/10/107,857
;  CURRENT FILING DATE:  2002-03-26
;  PRIOR APPLICATION NUMBER: 09/514,009
;  PRIOR FILING DATE: 2000-02-25
;  NUMBER OF SEQ ID NOS: 51
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 21
;   LENGTH: 795
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-107-857-21

  Query Match             72.9%;  Score 35;  DB 12;  Length 795;
  Best Local Similarity   75.0%;  Pred. No. 6.5e+02;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLQILWK 8
              ::||||||
Db        594 QQLQILWK 601


RESULT 15
US-10-107-857-45
; Sequence 45, Application US/10107857
; Publication No. US20030170784A1
; GENERAL INFORMATION:
;  APPLICANT: Fraser, Christopher C.
;  TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
;  FILE REFERENCE: 07334-251001
;  CURRENT APPLICATION NUMBER: US/10/107,857
;  CURRENT FILING DATE:  2002-03-26
;  PRIOR APPLICATION NUMBER: 09/514,009
;  PRIOR FILING DATE: 2000-02-25
;  NUMBER OF SEQ ID NOS: 51
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 45
;   LENGTH: 855
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-107-857-45

  Query Match             72.9%;  Score 35;  DB 12;  Length 855;
  Best Local Similarity   75.0%;  Pred. No. 6.9e+02;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLQILWK 8
              ::||||||
Db        654 QQLQILWK 661


Search completed: October 14, 2003, 14:44:51
Job time : 23.9375 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 14:33:18 ; Search time 11.8125 Seconds
(without alignments)
73.272 Million cell updates/sec

Title: US-09-881-490-163
Perfect score: 48
Sequence: 1 KKLQILWKK 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	39	81.2	339	1	R6BY13	ribosomal protein
2	39	81.2	1273	2	T38292	hypothetical prote
3	35	72.9	326	2	T24391	hypothetical prote
4	35	72.9	588	2	D70394	mannosyltransferas
5	35	72.9	795	2	T43447	hypothetical prote
6	35	72.9	867	2	JH0225	L96 protein - Tipu
7	35	72.9	2925	2	T00133	RNA-directed RNA p
8	34	70.8	137	2	D86439	protein T19E23.15
9	34	70.8	157	1	W6WLB5	E6 protein - human
10	34	70.8	205	2	B72277	phosphoribosylglyc
11	34	70.8	308	2	B86397	protein T7N9.8 [im
12	34	70.8	424	1	B64216	hemolysin tlyC hom
13	34	70.8	451	2	T19491	hypothetical prote
14	34	70.8	647	2	JC4239	phospholipase A2-a
15	34	70.8	1042	2	S76045	hypothetical prote
16	33	68.8	142	2	S36509	E6 protein - human
17	33	68.8	156	1	W6WL47	E6 protein - human
18	33	68.8	243	2	AD3405	transcription regu
19	33	68.8	263	2	G81272	hypothetical prote
20	33	68.8	313	2	A69496	TPR domain-contain
21	33	68.8	394	2	E81286	probable polysacch
22	33	68.8	412	2	B97077	diverged AAA-famil
23	33	68.8	413	2	T12533	hypothetical prote
24	33	68.8	476	2	C87454	PQQ enzyme repeat
25	33	68.8	502	1	A41361	serine/threonine-s
26	33	68.8	513	1	A47436	1,25-dihydroxyvita
27	33	68.8	513	1	S60303	serine/threonine-s
28	33	68.8	513	1	S60304	serine/threonine-s
29	33	68.8	514	2	H97819	hypothetical prote
30	33	68.8	541	2	T25258	hypothetical prote
31	33	68.8	563	2	T27313	hypothetical prote
32	33	68.8	578	2	A35810	alpha,alpha-trehal
33	33	68.8	610	2	T25262	hypothetical prote
34	33	68.8	625	2	AD2251	two-component sens
35	33	68.8	803	2	F59433	RhoGAP protein [im
36	33	68.8	818	2	A59433	KIAA0672 protein [
37	33	68.8	822	2	C72023	chltr probable pho
38	33	68.8	822	2	B81518	conserved hypothet
39	33	68.8	822	2	E86601	CHLTR probable pho
40	33	68.8	890	2	C96654	hypothetical prote
41	33	68.8	990	2	T16554	hypothetical prote
42	33	68.8	2628	2	S59413	probable membrane
43	33	68.8	3351	2	T13812	lipophorin - fruit
44	32.5	67.7	145	2	E70168	flagellar protein
45	32.5	67.7	665	2	E90351	ATP-dependent heli

ALIGNMENTS

RESULT 1
R6BY13
ribosomal protein MRP13 precursor, mitochondrial - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein G4618; protein YGR084c; ribosomal protein YmS-A
C;Species: Saccharomyces cerevisiae
C;Date: 30-Jun-1992 #sequence_revision 19-Jul-1996 #text_change 16-Jun-2000
C;Accession: S64379; A30237; S78035; S07831
R;Wedler, H.; Scharfe, M.; Wedler, E.; Wambutt, R.
submitted to the Protein Sequence Database, May 1996
A;Reference number: S64356
A;Accession: S64379
A;Molecule type: DNA
A;Residues: 1-339 <WED>
A;Cross-references: EMBL:Z72869; NID:g1323120; PIDN:CAA97086.1; PID:g1323121; GSPDB:G
A;Experimental source: strain S288C
R;Partaledis, J.A.; Mason, T.L.
Mol. Cell. Biol. 8, 3647-3660, 1988
A;Title: Structure and regulation of a nuclear gene in Saccharomyces cerevisiae that
A;Reference number: A30237; MUID:89127203; PMID:3065621
A;Accession: A30237
A;Molecule type: DNA
A;Residues: 28-113,'S',115-309,'RIQRSIRHAFESV',323,'TVGNTLSGSG',336,'RG',339,'IVQNTHR
A;Cross-references: EMBL:M22109; NID:g171970; PIDN:AAA34788.1; PID:g171971
A;Note: the authors translated the codon AGT for residues 4, 227, 236, and 282 as Phe
R;Kitakawa, M.; Graack, H.R.; Grohmann, L.; Goldschmidt-Reisin, S.; Herfurth, E.; Wit
Eur. J. Biochem. 245, 449-456, 1997
A;Title: Identification and characterization of the genes for mitochondrial ribosomal
A;Reference number: S78018; MUID:97296414; PMID:9151978
A;Accession: S78035
A;Molecule type: protein
A;Residues: 'X',39-51 <KIT>
C;Genetics:
A;Gene: SGD:MRP13; MIPS:YGR084c
A;Cross-references: SGD:S0003316; MIPS:YGR084c
A;Map position: 7R
A;Genome: nuclear
C;Superfamily: Saccharomyces cerevisiae mitochondrial ribosomal protein MRP13
C;Keywords: mitochondrion; protein biosynthesis; ribosome
F;1-47/Domain: transit peptide (mitochondrion) #status predicted <TNP>
F;48-339/Product: ribosomal protein MRP13 #status predicted <MAT>

Query Match 81.2%; Score 39; DB 1; Length 339;
Best Local Similarity 66.7%; Pred. No. 9.5;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 KKLQILWKK 9
              :|||: |||
Db        209 RKLQVAWKK 217
```

RESULT 2
T38292

```
hypothetical protein SPAC23E2.02 - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C;Accession: T38292
R;Skelton, J.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, January 1996
A;Reference number: Z21784
A;Accession: T38292
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1273 <SKE>
A;Cross-references: EMBL:Z68887; PIDN:CAA93114.1; GSPDB:GN00066; SPDB:SPAC23E2.02
A;Experimental source: strain 972h-; cosmid c23E2
C;Genetics:
A;Gene: SPDB:SPAC23E2.02
A;Map position: 1
A;Introns: 8/2; 862/3

  Query Match             81.2%;  Score 39;  DB 2;  Length 1273;
  Best Local Similarity   77.8%;  Pred. No. 35;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKLQILWKK 9
              ||| |||:|
Db        657 KKLNILWEK 665


RESULT 3
T24391
hypothetical protein T03E6.5 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C;Accession: T24391
R;Lloyd, C.
submitted to the EMBL Data Library, March 1997
A;Reference number: Z19885
A;Accession: T24391
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-326 <WIL>
A;Cross-references: EMBL:Z92812; PIDN:CAB07279.1; GSPDB:GN00023; CESP:T03E6.5
A;Experimental source: clone T03E6
C;Genetics:
A;Gene: CESP:T03E6.5
A;Map position: 5
A;Introns: 99/1; 160/3; 236/3
C;Superfamily: Caenorhabditis hypothetical protein C49G7.2

  Query Match             72.9%;  Score 35;  DB 2;  Length 326;
  Best Local Similarity   85.7%;  Pred. No. 51;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLQILW 7
              :||||||
Db        300 RKLQILW 306


RESULT 4
D70394
mannosyltransferase A - Aquifex aeolicus
C;Species: Aquifex aeolicus
C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999
C;Accession: D70394
R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov
V.
Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A;Reference number: A70300; MUID:98196666; PMID:9537320
A;Accession: D70394
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-588 <AQF>
A;Cross-references: GB:AE000723; NID:g2983569; PIDN:AAC07142.1; PID:g2983570; GB:AE0006
A;Experimental source: strain VF5
C;Genetics:
A;Gene: mtfA

  Query Match             72.9%;  Score 35;  DB 2;  Length 588;
  Best Local Similarity   66.7%;  Pred. No. 92;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 KKLQILWKK 9
              |  :|||||
Db        432 KAFEILWKK 440


RESULT 5
T43447
hypothetical protein DKFZp434J1217.1 - human
C;Species: Homo sapiens (man)
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 02-Sep-2000
C;Accession: T43447
R;Bloecker, H.; Boecher, M.; Brandt, P.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, December 1999
A;Reference number: Z22513
A;Accession: T43447
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-795 <AAA>
A;Cross-references: EMBL:AL133608
A;Experimental source: adult testis; clone DKFZp434J1217
C;Genetics:
A;Note: DKFZp434J1217.1
C;Superfamily: unassigned WD repeat proteins; WD repeat homology

  Query Match             72.9%;  Score 35;  DB 2;  Length 795;
  Best Local Similarity   75.0%;  Pred. No. 1.2e+02;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLQILWK 8
              ::||||||
Db        594 QQLQILWK 601


RESULT 6
JH0225
L96 protein - Tipula iridescent virus
C;Species: Tipula iridescent virus
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C;Accession: JH0225
R;Home, W.A.; Tajbakhsh, S.; Seligy, V.L.
Gene 94, 243-248, 1990
A;Title: Molecular cloning and characterization of a late Tipula iridescent virus gen
A;Reference number: JH0225; MUID:91078646; PMID:1701750
A;Accession: JH0225
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-867 <HOM>
A;Cross-references: GB:M34853; NID:g335215; PIDN:AAA47919.1; PID:g335216

  Query Match             72.9%;  Score 35;  DB 2;  Length 867;
  Best Local Similarity   77.8%;  Pred. No. 1.3e+02;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 KKLQILWKK 9
              ||| | |||
Db        152 KKLDIDWKK 160


RESULT 7
T00133
RNA-directed RNA polymerase (EC 2.7.7.48) - rice grassy stunt virus
C;Species: rice grassy stunt virus
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 21-Jul-2000
C;Accession: T00133
R;Toriyama, S.; Kimishima, T.; Takahashi, M.; Shimizu, T.; Minaka, N.; Akutsu, K.
```

J. Gen. Virol. 79, 2051-2058, 1998
A;Title: The complete nucleotide sequence of the rice grassy stunt virus genome and genc
A;Reference number: Z14118; MUID:98378059; PMID:9714257
A;Accession: T00133
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: genomic RNA
A;Residues: 1-2925 <TOR>
A;Cross-references: EMBL:AB009656; NID:g3410897; PIDN:BAA32246.1; PID:g3410899
A;Experimental source: isolate IRRI
C;Keywords: nucleotidyltransferase

 Query Match 72.9%; Score 35; DB 2; Length 2925;
 Best Local Similarity 66.7%; Pred. No. 4.5e-02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KKLQILWKK 9
| |:|:|||
Db 2209 KTLKIVWKK 2217

RESULT 8
D86439
protein T19E23.15 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 17-May-2002
C;Accession: D86439
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: D86439
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-137 <STO>
A;Cross-references: GB:AE005172; NID:g6692124; PIDN:AAF24589.1; GSPDB:GN00141
C;Genetics:
A;Gene: T19E23.15
A;Map position: 1
C;Superfamily: rat ribosomal protein L21

 Query Match 70.8%; Score 34; DB 2; Length 137;
 Best Local Similarity 71.4%; Pred. No. 33;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LQILWKK 9
|::||||
Db 5 LEVLWKK 11

RESULT 9
W6WLB5
E6 protein - human papillomavirus type 5b
C;Species: human papillomavirus type 5b
A;Note: host Homo sapiens (man)
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jun-2000
C;Accession: E40480
R;Yabe, Y.; Sakai, A.; Hitsumoto, T.; Kato, H.; Ogura, H.
Virology 183, 793-798, 1991
A;Title: A subtype of human papillomavirus 5 (HPV-5b) and its subgenomic segment amplifi
A;Reference number: A40480; MUID:91306467; PMID:1649510
A;Accession: E40480
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-157 <YAB>
A;Cross-references: GB:D90252; NID:g222395; PIDN:BAA14292.1; PID:g222398
C;Superfamily: papillomavirus E6 protein
C;Keywords: DNA binding; early protein; zinc finger
F;41-77/Region: zinc finger CCCC motif
F;114-150/Region: zinc finger CCCC motif

 Query Match 70.8%; Score 34; DB 1; Length 157;
 Best Local Similarity 62.5%; Pred. No. 38;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KKLQILWK 8
||| ::||
Db 59 KKLSLIWK 66

RESULT 10
B72277
phosphoribosylglycinamide formyltransferase - Thermotoga maritima (strain MSB8)
C;Species: Thermotoga maritima
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C;Accession: B72277
R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hi
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson,
C.M.
Nature 399, 323-329, 1999
A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome
A;Reference number: A72200; MUID:99287316; PMID:10360571
A;Accession: B72277
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-205 <ARN>
A;Cross-references: GB:AE001780; GB:AE000512; NID:g4981796; PIDN:AAD36323.1; PID:g49
A;Experimental source: strain MSB8
C;Genetics:
A;Gene: TM1248
C;Superfamily: phosphoribosylglycinamide formyltransferase; phosphoribosylglycinamid

 Query Match 70.8%; Score 34; DB 2; Length 205;
 Best Local Similarity 66.7%; Pred. No. 50;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KKLQILWKK 9
||||| |::
Db 56 KKLQIPWER 64

RESULT 11
B86397
protein T7N9.8 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: B86397
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alo
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzi
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tall
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: B86397
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-308 <STO>
A;Cross-references: GB:AE005172; NID:g8778853; PIDN:AAF79852.1; GSPDB:GN00141
C;Genetics:
A;Gene: T7N9.8
A;Map position: 1

 Query Match 70.8%; Score 34; DB 2; Length 308;
 Best Local Similarity 71.4%; Pred. No. 74;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 KKLQILW 7
              ||:|:||
Db         93 KKVQVLW 99
```

RESULT 12
B64216
hemolysin tlyC homolog - Mycoplasma genitalium
C;Species: Mycoplasma genitalium
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 07-Dec-1999
C;Accession: B64216
R;Fraser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.D.;
M.; Fuhrmann, J.; Nguyen, D.; Utterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick, J.
, C.A.; Venter, J.C.
Science 270, 397-403, 1995
A;Title: The minimal gene complement of Mycoplasma genitalium.
A;Reference number: A64200; MUID:96026346; PMID:7569993
A;Accession: B64216
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-424 <TIGR>
A;Cross-references: GB:U39694; GB:L43967; NID:g1045822; PID:g1045829; TIGR:MG146
A;Experimental source: strain G-37
C;Genetics:
A;Genetic code: SGC3
C;Superfamily: hemolysin homolog tlyC

 Query Match 70.8%; Score 34; DB 1; Length 424;
 Best Local Similarity 66.7%; Pred. No. 1e+02;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 KKLQILWKK 9
              ||: | |||
Db        207 KKVMIKWKK 215
```

RESULT 13
T19491
hypothetical protein C26E1.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000
C;Accession: T19491
R;Sims, M.
submitted to the EMBL Data Library, October 1996
A;Reference number: Z19130
A;Accession: T19491
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-451 <WIL>
A;Cross-references: EMBL:Z81040; PIDN:CAB02783.1; GSPDB:GN00023; CESP:C26E1.3
A;Experimental source: clone C26E1
C;Genetics:
A;Gene: CESP:C26E1.3
A;Map position: 5
A;Introns: 16/3; 51/2; 164/3; 198/3; 249/3; 303/2; 348/3; 385/3; 408/2
C;Superfamily: Caenorhabditis elegans hypothetical protein C26E1.3

 Query Match 70.8%; Score 34; DB 2; Length 451;
 Best Local Similarity 62.5%; Pred. No. 1.1e+02;
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 KKLQILWK 8
              |||::||:
Db        353 KKLELLWQ 360
```

RESULT 14
JC4239
phospholipase A2-activating protein - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 21-Nov-1995 #sequence_revision 08-Feb-1996 #text_change 26-May-2000
C;Accession: JC4239
R;Wang, H.; Lemasters, J.J.; Herman, B.
Gene 161, 237-241, 1995
A;Title: Cloning of a rat cDNA encoding a protein with high homology to mouse phospho
A;Reference number: JC4239; MUID:95394363; PMID:7665086
A;Accession: JC4239
A;Molecule type: mRNA
A;Residues: 1-647 <WAN>
A;Cross-references: GB:U17901
A;Experimental source: hepatocytes
A;Note: This protein has a homologous region, 503-538, with bee venom melittin
C;Comment: This protein is rich in cysteine, it contains fifteen Cys residues in the
C;Genetics:
A;Gene: hcw9
C;Superfamily: unassigned WD repeat proteins; WD repeat homology
F;90-123/Domain: WD repeat homology <WD1>
F;130-162/Domain: WD repeat homology <WD2>
F;170-203/Domain: WD repeat homology <WD3>

 Query Match 70.8%; Score 34; DB 2; Length 647;
 Best Local Similarity 62.5%; Pred. No. 1.6e+02;
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 KKLQILWK 8
              ::||:|||
Db        537 QQLQVLWK 544
```

RESULT 15
S76045
hypothetical protein - Synechocystis sp. (strain PCC 6803)
C;Species: Synechocystis sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C;Accession: S76045
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,
o, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas
DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocys
s.
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S76045
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1042 <KAN>
A;Cross-references: EMBL:D63999; GB:AB001339; NID:g1001396; PIDN:BAA10023.1; PID:g100
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C;Superfamily: Saccharomyces alpha-mannosidase

 Query Match 70.8%; Score 34; DB 2; Length 1042;
 Best Local Similarity 55.6%; Pred. No. 2.5e+02;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 KKLQILWKK 9
              :||:: |||
Db        545 EKLELAWKK 553
```

Search completed: October 14, 2003, 14:38:06
Job time : 14.8125 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 14:33:18 ; Search time 6.375 Seconds
(without alignments)
66.391 Million cell updates/sec

Title: US-09-881-490-163
Perfect score: 48
Sequence: 1 KKLQILWKK 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	39	81.2	339	1	RT13_YEAST	P12686 saccharomyc
2	39	81.2	1273	1	YAR2_SCHPO	Q10135 schizosacch
3	36	75.0	168	1	VE6_HPV2:	P28832 human papil
4	35	72.9	279	1	LACR_STAXY	O33813 staphylococ
5	35	72.9	646	1	PLAP_MOUSE	P27612 mus musculu
6	35	72.9	738	1	PLAP_HUMAN	Q9y263 homo sapien
7	35	72.9	740	1	NLD2_HUMAN	Q9y3q0 homo sapien
8	35	72.9	867	1	VL96_IRV1	P22856 tipula irid
9	34	70.8	157	1	VE6_HPV36	P50810 human papil
10	34	70.8	157	1	VE6_HPV5B	P26556 human papil
11	34	70.8	171	1	VE5_HPV14	P28830 human papil
12	34	70.8	424	1	Y146_MYCGE	Q49399 mycoplasma
13	34	70.8	647	1	PLAP_RAT	P54319 rattus norv
14	33	68.8	142	1	VE6_HPV32	P36810 human papil
15	33	68.8	156	1	VE6_HPV47	P22422 human papil
16	33	68.8	165	1	VE6_HPV20	P28831 human papil
17	33	68.8	493	1	OCLN_XENLA	Q9pun1 xenopus lae
18	33	68.8	502	1	RKI1_SECCE	Q02723 secale cere
19	33	68.8	514	1	CP24_HUMAN	Q07973 homo sapien
20	33	68.8	578	1	TREA_RABIT	P19813 oryctolagus
21	32	66.7	159	1	VE6_HPV27	P36808 human papil
22	32	66.7	161	1	VE6_HPV25	P28833 human papil
23	32	66.7	178	1	ARF3_CAEEL	Q94231 caenorhabdi
24	32	66.7	338	1	PHND_ECOLI	P16682 escherichia
25	32	66.7	452	1	RN18_HUMAN	Q9ns80 homo sapien
26	32	66.7	489	1	NIFD_ALCFA	Q44045 alcaligenes
27	32	66.7	518	1	DRT2_ARATH	Q05763 arabidopsis
28	32	66.7	535	1	DCTS_BACSU	P96601 bacillus su
29	32	66.7	645	1	REP_BUCAI	P57654 buchnera ap
30	31	64.6	157	1	VE6_HPV05	P06930 human papil
31	31	64.6	166	1	VE6_HPV19	P36806 human papil
32	31	64.6	185	1	SSRG_HUMAN	Q9un12 homo sapien
33	31	64.6	185	1	SSRG_RAT	Q08013 rattus norv
34	31	64.6	195	1	ENGB_THEMA	Q9x1h7 thermotoga
35	31	64.6	196	1	GTDB_AQUAE	O67522 aquifex aeo
36	31	64.6	214	1	ST14_SOLTU	Q41495 solanum tub
37	31	64.6	263	1	MT04_MYCTU	P96390 mycobacteri
38	31	64.6	283	1	YF46_ARCFU	O28726 archaeoglob
39	31	64.6	350	1	PE24_ARATH	Q9zv04 arabidopsis
40	31	64.6	416	1	SO_DROME	Q27350 drosophila
41	31	64.6	483	1	YC24_GUITH	O78473 guillardia
42	31	64.6	489	1	COAT_FMVD	P09519 figwort mos
43	31	64.6	519	1	TRPE_BUCDN	Q44697 buchnera ap
44	31	64.6	540	1	BB61_RABIT	Q05004 oryctolagus
45	31	64.6	580	1	P69_MYCHR	P15362 mycoplasma

ALIGNMENTS

RESULT 1
RT13_YEAST
ID RT13_YEAST STANDARD; PRT; 339 AA.
AC P12686;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE 40S ribosomal protein S13, mitochondrial precursor.
GN MRP13 OR YGR084C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE OF 28-339 FROM N.A.
RX MEDLINE=89127203; PubMed=3065621;
RA Partaledis J.A., Mason T.L.;
RT "Structure and regulation of a nuclear gene in Saccharomyces
RT cerevisiae that specifies MRP13, a protein of the small subunit of
RT the mitochondrial ribosome.";
RL Mol. Cell. Biol. 8:3647-3660(1988).
RN [2]
RP SEQUENCE FROM N.A.
RA Wedler H., Scharfe M., Wedler E., Wambutt R.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
CC -!- FUNCTION: COMPONENT OF THE SMALL SUBUNIT OF MITOCHONDRIAL
CC RIBOSOME.
CC -!- SUBCELLULAR LOCATION: Mitochondrial.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; Z72869; CAA97086.1; -.
DR EMBL; M22109; AAA34788.1; -.
DR PIR; S64379; R6BY13.
DR SGD; S0003316; MRP13.
KW Ribosomal protein; Mitochondrion; Transit peptide.
FT TRANSIT 1 37 MITOCHONDRION (POTENTIAL).
FT CHAIN 38 339 40S RIBOSOMAL PROTEIN S13.
FT CONFLICT 114 114 T -> S (IN REF. 1).
FT CONFLICT 310 339 ESSDPYDTHLKALGLLETPPPQESTEKEEK -> RIQRSIR
FT HAFESVGTVGNTLSGSGKRGKIVQNTHRKYINNIL (IN
FT REF. 1).
SQ SEQUENCE 339 AA; 38988 MW; 89DC5A71799F2CEE CRC64;

Query Match 81.2%; Score 39; DB 1; Length 339;
Best Local Similarity 66.7%; Pred. No. 3.5;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KKLQILWKK 9
:|||: |||

```
Db        209 RKLQVAWKK 217


RESULT 2
YAR2_SCHPO
ID   YAR2_SCHPO     STANDARD;      PRT;  1273 AA.
AC   Q10135;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Hypothetical protein C23E2.02 in chromosome I.
GN   SPAC23E2.02.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RX   MEDLINE=21848401; PubMed=11859360;
RA   Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA   Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA   Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA   Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA   Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA   Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA   James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA   Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA   Oliver K., O'Neil S., Pearson D., Quail M.A., Rabbinowitsch E.,
RA   Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA   Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA   Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA   Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA   Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA   Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
RA   Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA   Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA   Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA   Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA   Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA   Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA   Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA   Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA   Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT   "The genome sequence of Schizosaccharomyces pombe.";
RL   Nature 415:871-880(2002).
CC   -!- SIMILARITY: Contains 1 HMG box domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z68887; CAA93114.1; -.
DR   PIR; T38292; T38292.
DR   GeneDB_SPombe; SPAC23E2.02; -.
DR   InterPro; IPR002937; Amino_oxidase.
DR   InterPro; IPR000910; HMG_12_box.
DR   Pfam; PF01593; Amino_oxidase; 1.
DR   Pfam; PF04433; SWIRM; 1.
DR   PROSITE; PS50118; HMG_BOX_2; 1.
KW   Hypothetical protein; Transmembrane.
FT   TRANSMEM    1028   1048       POTENTIAL.
FT   DNA_BIND    1115   1195       HMG BOX.
SQ   SEQUENCE   1273 AA;  142489 MW;  3FEFF6433DF3C0B6 CRC64;

  Query Match             81.2%;  Score 39;  DB 1;  Length 1273;
  Best Local Similarity   77.8%;  Pred. No. 13;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKLQILWKK 9
              ||| |||:|
Db        657 KKLNILWEK 665


RESULT 3
VE6_HPV21
ID   VE6_HPV21      STANDARD;      PRT;   168 AA.
AC   P28832;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   E6 protein.
GN   E6.
OS   Human papillomavirus type 21.
OC   Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC   Papillomavirus.
OX   NCBI_TaxID=31548;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92124737; PubMed=1310189;
RA   Kiyono T., Hiraiwa A., Ishibashi M.;
RT   "Differences in transforming activity and coded amino acid sequence
RT   among E6 genes of several papillomaviruses associated with
RT   epidermodysplasia verruciformis.";
RL   Virology 186:628-639(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Delius H.;
RL   Submitted (OCT-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: EXHIBIT A STRONG, BUT NON SPECIFIC AFFINITY FOR DOUBLE
CC       STRANDED DNA (IN VITRO).
CC   -!- SUBCELLULAR LOCATION: Nuclear matrix-associated.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D90263; BAA14310.1; -.
DR   EMBL; U31779; AAA79394.1; -.
DR   InterPro; IPR001334; E6.
DR   Pfam; PF00518; E6; 1.
KW   Early protein; DNA-binding; Nuclear protein; Zinc-finger.
FT   ZN_FING      55     91       POTENTIAL.
FT   ZN_FING     128    164       POTENTIAL.
FT   CONFLICT    130    130       T -> N (IN REF. 2).
FT   CONFLICT    148    148       H -> F (IN REF. 2).
FT   CONFLICT    156    156       T -> S (IN REF. 2).
SQ   SEQUENCE   168 AA;  19290 MW;  744E8AF8F0ACE53C CRC64;

  Query Match             75.0%;  Score 36;  DB 1;  Length 168;
  Best Local Similarity   75.0%;  Pred. No. 6.3;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKLQILWK 8
              ||| :|||
Db         73 KKLSLLWK 80


RESULT 4
LACR_STAXY
ID   LACR_STAXY     STANDARD;      PRT;   279 AA.
AC   O33813;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   Lactose operon transcription activator.
GN   LACR.
```

```
OS   Staphylococcus xylosus.
OC   Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX   NCBI_TaxID=1288;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DSM 20267 / Isolate C2A;
RX   MEDLINE=98233718; PubMed=9573174;
RA   Bassias J., Brueckner R.;
RT   "Regulation of lactose utilization genes in Staphylococcus xylosus.";
RL   J. Bacteriol. 180:2273-2279(1998).
CC   -!- FUNCTION: TRANSCRIPTIONAL REGULATOR OF THE LACPII GENES FOR
CC       LACTOSE UTILIZATION.
CC   -!- SIMILARITY: BELONGS TO THE ARAC/XYLS FAMILY OF TRANSCRIPTIONAL
CC       REGULATORS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; Y14599; CAA74935.1; -.
DR   InterPro; IPR000005; HTHAraC.
DR   Pfam; PF00165; HTH_AraC; 2.
DR   SMART; SM00342; HTH_ARAC; 1.
DR   PROSITE; PS00041; HTH_ARAC_FAMILY_1; 1.
DR   PROSITE; PS01124; HTH_ARAC_FAMILY_2; 1.
KW   Transcription regulation; DNA-binding; Activator.
FT   DNA_BIND    190    209       H-T-H MOTIF (BY SIMILARITY).
SQ   SEQUENCE   279 AA;  32335 MW;  304CDF65A7B14D34 CRC64;

  Query Match             72.9%;  Score 35;  DB 1;  Length 279;
  Best Local Similarity   71.4%;  Pred. No. 16;
  Matches    5;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LQILWKK 9
              :|:||||
Db          1 MQVLWKK 7


RESULT 5
PLAP_MOUSE
ID   PLAP_MOUSE     STANDARD;      PRT;   646 AA.
AC   P27612;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Phospholipase A-2-activating protein (PLAP).
GN   PLAA OR PLAP.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91271400; PubMed=2052621;
RA   Clark M.A., Oezguer L.E., Conway T.M., Dispoto J., Crooke S.T.,
RA   Bomalaski J.S.;
RT   "Cloning of a phospholipase A2-activating protein.";
RL   Proc. Natl. Acad. Sci. U.S.A. 88:5418-5422(1991).
RN   [2]
RP   REVISIONS.
RX   MEDLINE=95394363; PubMed=7665086;
RA   Wang H., Lemasters J.J., Herman B.;
RT   "Cloning of a rat cDNA encoding a protein with high homology to mouse
RT   phospholipase A2-activating protein.";
RL   Gene 161:237-241(1995).
CC   -!- FUNCTION: PLAY AN IMPORTANT ROLE IN THE REGULATION OF SPECIFIC
CC       INFLAMMATORY DISEASE PROCESSES.
CC   -!- INDUCTION: IN SMOOTH MUSCLE & ENDOTHELIAL CELLS BY LEUKOTRIENE D4,
CC       BY TUMOR NECROSIS FACTOR IN ENDOTHELIAL CELLS AND BY URIC ACID
CC       CRYSTALS IN MACROPHAGES.
CC   -!- SIMILARITY: Contains 6 WD repeats.
CC   -!- SIMILARITY: BELONGS TO THE PLAP FAMILY OF WD-REPEAT PROTEINS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M57958; AAA39943.1; ALT_SEQ.
DR   MGD; MGI:104810; Plaa.
DR   GO; GO:0016005; F:phospholipase A2 activator activity; IDA.
DR   GO; GO:0005515; F:protein binding activity; IPI.
DR   GO; GO:0006954; P:inflammatory response; IDA.
DR   InterPro; IPR001680; WD40.
DR   Pfam; PF00400; WD40; 6.
DR   PRINTS; PR00320; GPROTEINBRPT.
DR   ProDom; PD000018; WD40; 1.
DR   SMART; SM00320; WD40; 5.
DR   PROSITE; PS00678; WD_REPEATS_1; FALSE_NEG.
DR   PROSITE; PS50082; WD_REPEATS_2; 3.
DR   PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW   Repeat; WD repeat.
FT   REPEAT       6     41       WD 1.
FT   REPEAT      53     82       WD 2.
FT   REPEAT      92    122       WD 3.
FT   REPEAT     132    161       WD 4.
FT   REPEAT     172    202       WD 5.
FT   REPEAT     212    241       WD 6.
SQ   SEQUENCE   646 AA;  71416 MW;  8219C118DBE080C5 CRC64;

  Query Match             72.9%;  Score 35;  DB 1;  Length 646;
  Best Local Similarity   75.0%;  Pred. No. 39;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLQILWK 8
              ::||||||
Db        537 QQLQILWK 544


RESULT 6
PLAP_HUMAN
ID   PLAP_HUMAN     STANDARD;      PRT;   738 AA.
AC   Q9Y263; Q9Y5L1;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Phospholipase A-2-activating protein (PLAP) (PLA2P).
GN   PLAA OR PLAP.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fetal brain;
RA   Ruiz A., Nadal M., Puig S., Estivill X.;
RT   "Cloning of the human phospholipase A2 activating protein (hPLAP) gene
RT   on the chromosome 9p21 melanoma deleted region.";
RL   Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99132024; PubMed=9931468;
RA   Chopra A.K., Ribardo D.A., Wood T.G., Prusak D.J., Xu X.-J.,
RA   Peterson J.W.;
RT   "Molecular characterization of cDNA for phospholipase A2-activating
RT   protein.";
RL   Biochim. Biophys. Acta 1444:125-130(1999).
RN   [3]
RP   SEQUENCE FROM N.A.
```

```
RX   MEDLINE=20112769; PubMed=10644453;
RA   Beatty B., Qi S., Pienkowska M., Scherer S.W., Testa J.R., Cheng J.Q.,
RA   Herbrick J.-A., Scheidl T., Zhang Z., Kola I., Seth A.;
RT   "Chromosomal localization of phospholipase A2 activating protein, an
RT   ets2 target gene, to 9p21.";
RL   Genomics 62:529-532(1999).
CC   -!- FUNCTION: PLAY AN IMPORTANT ROLE IN THE REGULATION OF SPECIFIC
CC       INFLAMMATORY DISEASE PROCESSES (BY SIMILARITY).
CC   -!- SIMILARITY: Contains 6 WD repeats.
CC   -!- SIMILARITY: BELONGS TO THE PLAP FAMILY OF WD-REPEAT PROTEINS.
CC   -!- CAUTION: REF.3 SEQUENCE DIFFERS FROM THAT SHOWN IN POSITION 642
CC       ONWARD DUE TO A FRAMESHIFT.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AJ238243; CAB42881.1; -.
DR   EMBL; AF083395; AAD03030.1; -.
DR   EMBL; AF145020; AAD42075.1; ALT_FRAME.
DR   Genew; HGNC:9043; PLAA.
DR   MIM; 603873; -.
DR   GO; GO:0016005; F:phospholipase A2 activator activity; TAS.
DR   GO; GO:0006644; P:phospholipid metabolism; TAS.
DR   GO; GO:0007165; P:signal transduction; TAS.
DR   InterPro; IPR001680; WD40.
DR   Pfam; PF00400; WD40; 6.
DR   PRINTS; PR00320; GPROTEINBRPT.
DR   ProDom; PD000018; WD40; 1.
DR   SMART; SM00320; WD40; 6.
DR   PROSITE; PS00678; WD_REPEATS_1; FALSE_NEG.
DR   PROSITE; PS50082; WD_REPEATS_2; 3.
DR   PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW   Repeat; WD repeat.
FT   REPEAT        6     50       WD 1.
FT   REPEAT       53     91       WD 2.
FT   REPEAT       92    131       WD 3.
FT   REPEAT      133    170       WD 4.
FT   REPEAT      172    211       WD 5.
FT   REPEAT      213    250       WD 6.
FT   CONFLICT     29     29       A -> P (IN REF. 3).
FT   CONFLICT     40     40       P -> L (IN REF. 3).
FT   CONFLICT    365    365       I -> A (IN REF. 3).
FT   CONFLICT    474    474       M -> L (IN REF. 3).
FT   CONFLICT    484    484       V -> L (IN REF. 3).
SQ   SEQUENCE   738 AA;  80826 MW;  F2DB210E5C9B381A CRC64;

  Query Match             72.9%;  Score 35;  DB 1;  Length 738;
  Best Local Similarity   75.0%;  Pred. No. 44;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLQILWK 8
              ::|||||
Db        537 QQLQILWK 544


RESULT 7
NLD2_HUMAN
ID   NLD2_HUMAN      STANDARD;      PRT;   740 AA.
AC   Q9Y3Q0;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   N-acetylated-alpha-linked acidic dipeptidase II (EC 3.4.17.21)
DE   (NAALADase II).
GN   NAALAD2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Lung carcinoma;
RX   MEDLINE=99185063; PubMed=10085079;
RA   Pangalos M.N., Neefs J.-M., Somers M., Verhasselt P., Bekkers M.,
RA   van der Helm L., Fraiponts E., Ashton D., Gordon R.D.;
RT   "Isolation and expression of novel human glutamate carboxypeptidases
RT   with N-acetylated alpha-linked acidic dipeptidase and dipeptidyl
RT   peptidase IV activity.";
RL   J. Biol. Chem. 274:8470-8483(1999).
CC   -!- FUNCTION: Has N-acetylated-alpha-linked-acidic dipeptidase
CC       (NAALADase) activity. Also exhibits a dipeptidyl-peptidase IV type
CC       activity. In vitro, cleaves Gly-Pro-AMC.
CC   -!- CATALYTIC ACTIVITY: Release of an unsubstituted, C-terminal
CC       glutamyl residue, typically from Ac-Asp-Glu or pteroyl-gamma.
CC   -!- COFACTOR: Zinc; Binds two ions per subunit. Required for NAALADase
CC       activity.
CC   -!- ENZYME REGULATION: Inhibited by quisqualate.
CC   -!- SUBCELLULAR LOCATION: Type II membrane protein (Potential).
CC   -!- TISSUE SPECIFICITY: Highest expression in the testis. Also found
CC       in ovary and spleen. Weak expression in prostate, heart and
CC       placenta. In brain, expressed in striatum, parietal cortex and
CC       ventral striatum with lower levels in hippocampus, brain stem,
CC       putamen and superior colliculus.
CC   -!- DOMAIN: The NAALADase activity is found in the central region, the
CC       dipeptidyl peptidase IV type activity in the C-terminal.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M28B.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AJ012370; CAB39967.1; -.
DR   MEROPS; M28.012; -.
DR   Genew; HGNC:14526; NAALAD2.
DR   GO; GO:0016021; C:integral to membrane; NAS.
DR   GO; GO:0004180; F:carboxypeptidase activity; NAS.
DR   GO; GO:0016805; F:dipeptidase activity; NAS.
DR   GO; GO:0004274; F:dipeptidyl-peptidase IV activity; NAS.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; NAS.
DR   InterPro; IPR003137; PA.
DR   Pfam; PF02225; PA; 1.
DR   Pfam; PF04389; Peptidase_M28; 1.
DR   Pfam; PF04253; TFR_dimer; 1.
KW   Hydrolase; Carboxypeptidase; Metalloprotease; Zinc; Aminopeptidase;
KW   Dipeptidase; Serine protease; Transmembrane; Glycoprotein;
KW   Signal-anchor; Multifunctional enzyme.
FT   DOMAIN        1      7       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM      8     31       SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT                                (POTENTIAL).
FT   DOMAIN       32    740       EXTRACELLULAR (POTENTIAL).
FT   DOMAIN      264    577       NAALADASE.
FT   ACT_SITE    414    414       NUCLEOPHILE (NAALADASE) (BY SIMILARITY).
FT   ACT_SITE    618    618       CHARGE RELAY SYSTEM (POTENTIAL).
FT   ACT_SITE    656    656       CHARGE RELAY SYSTEM (POTENTIAL).
FT   ACT_SITE    679    679       CHARGE RELAY SYSTEM (POTENTIAL).
FT   METAL       367    367       ZINC 2 (BY SIMILARITY).
FT   METAL       377    377       ZINC 1 AND 2 (BY SIMILARITY).
FT   METAL       415    415       ZINC 1 (BY SIMILARITY).
FT   METAL       443    443       ZINC 2 (BY SIMILARITY).
FT   METAL       543    543       ZINC 1 (BY SIMILARITY).
FT   CARBOHYD    111    111       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    143    143       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    185    185       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    314    314       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    449    449       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    603    603       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    628    628       N-LINKED (GLCNAC. . .) (POTENTIAL).
```

```
SQ   SEQUENCE   740 AA;  83591 MW;  040624D691ECF879 CRC64;

  Query Match             72.9%;  Score 35;  DB 1;  Length 740;
  Best Local Similarity   66.7%;  Pred. No. 44;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 KKLQILWKK 9
              ||:|  ||| 
Db         84 KKIQTQWKK 92


RESULT 8
VL96_IRV1
ID   VL96_IRV1      STANDARD;      PRT;   867 AA.
AC   P22856;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   L96 protein.
GN   L96.
OS   Tipula iridescent virus (TIV) (Insect iridescent virus type 1).
OC   Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Iridovirus.
OX   NCBI_TaxID=10490;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91078646; PubMed=1701750;
RA   Home W.A., Tajbakhsh S., Seligy V.L.;
RT   "Molecular cloning and characterization of a late Tipula iridescent
RT   virus gene.";
RL   Gene 94:243-248(1990).
CC   -!- FUNCTION: MAY BE INVOLVED IN TIV GENOMIC DNA PACKAGING IN A
CC       MANNER RELATED TO THE GAG POLYPROTEINS OF THE MAMMALIAN VIRUSES.
CC   -!- SIMILARITY: Contains 1 OTU domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M62953; AAA47919.1; -.
DR   PIR; JH0225; JH0225.
DR   InterPro; IPR003323; OTU.
DR   Pfam; PF02338; OTU; 1.
DR   PROSITE; PS50802; OTU; 1.
KW   Repeat; DNA packaging; DNA-binding.
FT   DOMAIN      606    745       OTU.
SQ   SEQUENCE   867 AA;  96011 MW;  F19DBDE8FE5CA103 CRC64;

  Query Match             72.9%;  Score 35;  DB 1;  Length 867;
  Best Local Similarity   77.8%;  Pred. No. 52;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 KKLQILWKK 9
              ||| : ||| 
Db        152 KKLDIDWKK 160


RESULT 9
VE6_HPV36
ID   VE6_HPV36      STANDARD;      PRT;   157 AA.
AC   P50810;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   E6 protein.
GN   E6.
OS   Human papillomavirus type 36.
OC   Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC   Papillomavirus.
OX   NCBI_TaxID=37957;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Delius H.;
RL   Submitted (OCT-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: EXHIBIT A STRONG, BUT NON SPECIFIC AFFINITY FOR DOUBLE
CC       STRANDED DNA (IN VITRO).
CC   -!- SUBCELLULAR LOCATION: Nuclear matrix-associated.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U31785; AAA79436.1; -.
DR   InterPro; IPR001334; E6.
DR   Pfam; PF00518; E6; 1.
KW   Early protein; DNA-binding; Nuclear protein; Zinc-finger.
FT   ZN_FING      41     77       POTENTIAL.
FT   ZN_FING     114    150       POTENTIAL.
SQ   SEQUENCE   157 AA;  18123 MW;  ECB6CE02C0ECB126 CRC64;

  Query Match             70.8%;  Score 34;  DB 1;  Length 157;
  Best Local Similarity   62.5%;  Pred. No. 14;
  Matches    5;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKLQILWK 8
              ||| ::||
Db         59 KKLSLIWK 66


RESULT 10
VE6_HPV5B
ID   VE6_HPV5B      STANDARD;      PRT;   157 AA.
AC   P26556;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   E6 protein.
GN   E6.
OS   Human papillomavirus type 5b.
OC   Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC   Papillomavirus.
OX   NCBI_TaxID=10599;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91306467; PubMed=1649510;
RA   Yabe Y., Sakai A., Hitsumoto T., Kato H., Ogura H.;
RT   "A subtype of human papillomavirus 5 (HPV-5b) and its subgenomic
RT   segment amplified in a carcinoma: nucleotide sequences and genomic
RT   organizations.";
RL   Virology 183:793-798(1991).
CC   -!- FUNCTION: EXHIBIT A STRONG, BUT NON SPECIFIC AFFINITY FOR DOUBLE
CC       STRANDED DNA (IN VITRO).
CC   -!- SUBCELLULAR LOCATION: Nuclear matrix-associated.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D90252; BAA14292.1; -.
DR   PIR; E40480; W6WLB5.
DR   InterPro; IPR001334; E6.
DR   Pfam; PF00518; E6; 1.
KW   Early protein; DNA-binding; Nuclear protein; Zinc-finger.
FT   ZN_FING      41     77       POTENTIAL.
FT   ZN_FING     114    150       POTENTIAL.
```

```
SQ   SEQUENCE   157 AA;  18087 MW;  20684B7EF730AA47 CRC64;

  Query Match             70.8%;  Score 34;  DB 1;  Length 157;
  Best Local Similarity   62.5%;  Pred. No. 14;
  Matches    5;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKLQILWK 8
              ||| ::||
Db         59 KKLSLIWK 66


RESULT 11
VE6_HPV14
ID   VE6_HPV14      STANDARD;      PRT;   171 AA.
AC   P28830;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   E6 protein.
GN   E6.
OS   Human papillomavirus type 14.
OC   Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC   Papillomavirus.
OX   NCBI_TaxID=10605;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92124737; PubMed=1310189;
RA   Kiyono T., Hiraiwa A., Ishibashi M.;
RT   "Differences in transforming activity and coded amino acid sequence
RT   among E6 genes of several papillomaviruses associated with
RT   epidermodysplasia verruciformis.";
RL   Virology 186:628-639(1992).
CC   -!- FUNCTION: EXHIBIT A STRONG, BUT NON SPECIFIC AFFINITY FOR DOUBLE
CC       STRANDED DNA (IN VITRO).
CC   -!- SUBCELLULAR LOCATION: Nuclear matrix-associated.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D90262; BAA14309.1; -.
DR   InterPro; IPR001334; E6.
DR   Pfam; PF00518; E6; 1.
KW   Early protein; DNA-binding; Nuclear protein; Zinc-finger.
FT   ZN_FING      58     94       POTENTIAL.
FT   ZN_FING     131    167       POTENTIAL.
SQ   SEQUENCE   171 AA;  19214 MW;  394FBD3E5903F69F CRC64;

  Query Match             70.8%;  Score 34;  DB 1;  Length 171;
  Best Local Similarity   62.5%;  Pred. No. 16;
  Matches    5;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKLQILWK 8
              ||| ::||
Db         76 KKLSLIWK 83


RESULT 12
Y146_MYCGE
ID   Y146_MYCGE     STANDARD;      PRT;   424 AA.
AC   Q49399;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical protein MG146 (VXpSPT7_orf424).
GN   MG146 OR MPN159 OR MP672.
OS   Mycoplasma genitalium.
OC   Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX   NCBI_TaxID=2097;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 33530 / G-37;
RX   MEDLINE=96026346; PubMed=7569993;
RA   Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA   Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA   Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,
RA   Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA   Tomb J.-F., Dougherty B.A., Bott K.F., Hu P.-C., Lucier T.S.,
RA   Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RT   "The minimal gene complement of Mycoplasma genitalium.";
RL   Science 270:397-403(1995).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC   -!- SIMILARITY: BELONGS TO THE UPF0053 FAMILY.
CC   -!- SIMILARITY: Contains 2 CBS domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U39695; AAC71364.1; -.
DR   PIR; B64216; B64216.
DR   TIGR; MG146; -.
DR   InterPro; IPR000644; CBS_domain.
DR   InterPro; IPR002550; DUF21.
DR   Pfam; PF00571; CBS; 2.
DR   Pfam; PF01595; DUF21; 1.
DR   SMART; SM00116; CBS; 2.
KW   Hypothetical protein; CBS domain; Repeat; Transmembrane;
KW   Complete proteome.
FT   TRANSMEM      7     27       POTENTIAL.
FT   TRANSMEM     71     91       POTENTIAL.
FT   TRANSMEM    101    121       POTENTIAL.
FT   TRANSMEM    135    155       POTENTIAL.
SQ   SEQUENCE   424 AA;  48720 MW;  A361AAE5B995DA87 CRC64;

  Query Match             70.8%;  Score 34;  DB 1;  Length 424;
  Best Local Similarity   66.7%;  Pred. No. 39;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 KKLQILWKK 9
              ||: | |||
Db        207 KKVMIKWKK 215


RESULT 13
PLAP_RAT
ID   PLAP_RAT       STANDARD;      PRT;   647 AA.
AC   P54319;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Phospholipase A-2-activating protein (PLAP).
GN   PLAA OR PLAP.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Liver;
RX   MEDLINE=95394363; PubMed=7665086;
RA   Wang H., Lemasters J.J., Herman B.;
RT   "Cloning of a rat cDNA encoding a protein with high homology to mouse
RT   phospholipase A2-activating protein.";
RL   Gene 161:237-241(1995).
CC   -!- FUNCTION: PLAY AN IMPORTANT ROLE IN THE REGULATION OF SPECIFIC
CC       INFLAMMATORY DISEASE PROCESSES.
CC   -!- SIMILARITY: Contains 6 WD repeats.
```

```
CC   -!- SIMILARITY: BELONGS TO THE PLAP FAMILY OF WD-REPEAT PROTEINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U17901; AAA79979.1; ALT_SEQ.
DR   PIR; JC4239; JC4239.
DR   InterPro; IPR001680; WD40.
DR   Pfam; PF00400; WD40; 6.
DR   PRINTS; PR00320; GPROTEINBRPT.
DR   ProDom; PD000018; WD40; 1.
DR   SMART; SM00320; WD40; 6.
DR   PROSITE; PS00678; WD_REPEATS_1; FALSE_NEG.
DR   PROSITE; PS50082; WD_REPEATS_2; 3.
DR   PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW   Repeat; WD repeat.
FT   REPEAT        6     41       WD 1.
FT   REPEAT       53     82       WD 2.
FT   REPEAT       92    122       WD 3.
FT   REPEAT      132    161       WD 4.
FT   REPEAT      172    202       WD 5.
FT   REPEAT      212    241       WD 6.
SQ   SEQUENCE   647 AA;  71187 MW;  B0A8BD230AC5300C CRC64;

  Query Match             70.8%;  Score 34;  DB 1;  Length 647;
  Best Local Similarity   62.5%;  Pred. No. 60;
  Matches    5;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLQILWK 8
              ::||:|||
Db        537 QQLQVLWK 544


RESULT 14
VE6_HPV32
ID   VE6_HPV32      STANDARD;      PRT;   142 AA.
AC   P36810;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-OCT-1994 (Rel. 30, Last annotation update)
DE   E6 protein.
GN   E6.
OS   Human papillomavirus type 32.
OC   Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC   Papillomavirus.
OX   NCBI_TaxID=10612;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94265501; PubMed=8205838;
RA   Delius H., Hofmann B.;
RT   "Primer-directed sequencing of human papillomavirus types.";
RL   Curr. Top. Microbiol. Immunol. 186:13-31(1994).
CC   -!- FUNCTION: EXHIBIT A STRONG, BUT NON SPECIFIC AFFINITY FOR DOUBLE
CC       STRANDED DNA (IN VITRO).
CC   -!- SUBCELLULAR LOCATION: Nuclear matrix-associated.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X74475; CAA52549.1; -.
DR   PIR; S36509; S36509.
DR   InterPro; IPR001334; E6.
DR   Pfam; PF00518; E6; 1.
DR   SMART; SM00355; ZnF_C2H2; 1.
KW   Early protein; DNA-binding; Nuclear protein; Zinc-finger.
FT   ZN_FING      31     67       POTENTIAL.
FT   ZN_FING     104    140       POTENTIAL.
SQ   SEQUENCE   142 AA;  16631 MW;  7E108DDE4A72F297 CRC64;

  Query Match             68.8%;  Score 33;  DB 1;  Length 142;
  Best Local Similarity   55.6%;  Pred. No. 20;
  Matches    5;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 KKLQILWKK 9
              | | ::|||
Db         49 KDLHVVWKK 57


RESULT 15
VE6_HPV47
ID   VE6_HPV47      STANDARD;      PRT;   156 AA.
AC   P22422;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   E6 protein.
GN   E6.
OS   Human papillomavirus type 47.
OC   Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC   Papillomavirus.
OX   NCBI_TaxID=10594;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90281611; PubMed=2162112;
RA   Kiyono T., Adachi A., Ishibashi M.;
RT   "Genome organization and taxonomic position of human papillomavirus
RT   type 47 inferred from its DNA sequence.";
RL   Virology 177:401-405(1990).
CC   -!- FUNCTION: EXHIBIT A STRONG, BUT NON SPECIFIC AFFINITY FOR DOUBLE
CC       STRANDED DNA (IN VITRO).
CC   -!- SUBCELLULAR LOCATION: Nuclear matrix-associated.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M32305; AAA46976.1; -.
DR   PIR; A35324; W6WL47.
DR   InterPro; IPR001334; E6.
DR   Pfam; PF00518; E6; 1.
KW   Early protein; DNA-binding; Nuclear protein; Zinc-finger.
FT   ZN_FING      40     76       POTENTIAL.
FT   ZN_FING     113    149       POTENTIAL.
SQ   SEQUENCE   156 AA;  18116 MW;  A2AF8D1B95BCE623 CRC64;

  Query Match             68.8%;  Score 33;  DB 1;  Length 156;
  Best Local Similarity   62.5%;  Pred. No. 22;
  Matches    5;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKLQILWK 8
              ||| ::||
Db         58 KKLTLIWK 65


Search completed: October 14, 2003, 14:34:08
Job time : 8.54167 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 14:33:18 ; Search time 29.8125 Seconds
(without alignments)
77.903 Million cell updates/sec

Title: US-09-881-490-163
Perfect score: 48
Sequence: 1 KKLQILWKK 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	40	83.3	366	8	Q8HUW9	Q8huw9 ulmus ameri
2	39	81.2	701	11	Q8BQZ3	Q8bqz3 mus musculu
3	39	81.2	731	11	Q9JIE5	Q9jie5 mus musculu
4	39	81.2	745	11	Q9JI20	Q9ji20 mus musculu
5	39	81.2	745	11	Q8CC51	Q8cc51 mus musculu
6	39	81.2	746	11	P97616	P97616 rattus norv
7	38	79.2	446	4	Q96BQ3	Q96bq3 homo sapien
8	38	79.2	564	10	Q9LSS5	Q9lss5 arabidopsis
9	38	79.2	739	4	Q9NS39	Q9ns39 homo sapien
10	38	79.2	900	5	Q8I634	Q8i634 plasmodium
11	37	77.1	148	12	Q8JNA1	Q8jna1 human papil
12	36	75.0	208	4	Q9BUW4	Q9buw4 homo sapien
13	36	75.0	279	16	Q8XMB9	Q8xmb9 clostridium
14	36	75.0	375	4	Q8IWZ4	Q8iwz4 homo sapien
15	36	75.0	716	11	Q8K5B1	Q8k5b1 mus musculu
16	36	75.0	716	11	Q8R501	Q8r501 mus musculu
17	36	75.0	822	5	Q9BMX6	Q9bmx6 aplysia cal
18	36	75.0	1102	5	Q8IRF5	Q8irf5 drosophila
19	36	75.0	2217	13	Q8AV57	Q8av57 gallus gall
20	35	72.9	293	10	Q9FWJ9	Q9fwj9 arabidopsis
21	35	72.9	317	4	Q8IUX3	Q8iux3 homo sapien
22	35	72.9	326	5	O45738	O45738 caenorhabdi
23	35	72.9	357	8	Q8HUX0	Q8hux0 ampelocera
24	35	72.9	588	16	O67183	O67183 aquifex aeo
25	35	72.9	718	10	O82779	O82779 adiantum ca
26	35	72.9	794	11	Q8C6C4	Q8c6c4 mus musculu
27	35	72.9	795	4	Q9NVE9	Q9nve9 homo sapien
28	35	72.9	795	4	Q9UF53	Q9uf53 homo sapien
29	35	72.9	895	11	Q8C117	Q8c117 mus musculu
30	35	72.9	1043	16	Q9CMV5	Q9cmv5 pasteurella
31	35	72.9	2241	12	Q91HJ2	Q91hj2 avian param
32	35	72.9	2925	12	O89163	O89163 rice grassy
33	35	72.9	2925	12	Q9JGN8	Q9jgn8 rice grassy
34	34	70.8	62	12	Q91GK8	Q91gk8 epiphyas po
35	34	70.8	137	10	Q9SHE5	Q9she5 arabidopsis
36	34	70.8	157	12	Q913W2	Q913w2 human papil
37	34	70.8	157	12	Q913W4	Q913w4 human papil
38	34	70.8	157	12	Q913W5	Q913w5 human papil
39	34	70.8	157	12	Q81962	Q81962 human papil
40	34	70.8	157	12	Q913V9	Q913v9 human papil
41	34	70.8	157	12	Q913W0	Q913w0 human papil
42	34	70.8	157	12	Q913W6	Q913w6 human papil
43	34	70.8	157	12	Q913V8	Q913v8 human papil
44	34	70.8	157	12	Q913V7	Q913v7 human papil
45	34	70.8	157	12	Q81985	Q81985 human papil

ALIGNMENTS

RESULT 1
Q8HUW9
ID Q8HUW9 PRELIMINARY; PRT; 366 AA.
AC Q8HUW9;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE NADH dehydrogenase (Fragment).
GN NDHF.
OS Ulmus americana.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Rosales; Ulmaceae; Ulmus.
OX NCBI_TaxID=29740;
RN [1]
RP SEQUENCE FROM N.A.
RA Sytsma K.J., Morawetz J., Pires J.C., Nepokroeff M., Conti E.,
RA Zjhra M., Hall J.C., Chase M.W.;
RT "Urticalean rosids: circumscription, rosid ancestry, and phylogenetics
RT based on rbcL, trnL-F, and ndhF sequences.";
RL Am. J. Bot. 89:1531-1546(2002).
DR EMBL; AF500365; AAN63301.1; -.
KW Chloroplast.
FT NON_TER 1 1
FT NON_TER 366 366
SQ SEQUENCE 366 AA; 41953 MW; 4A74694C0B1073EC CRC64;

Query Match 83.3%; Score 40; DB 8; Length 366;
Best Local Similarity 77.8%; Pred. No. 12;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 KKLQILWKK 9
              ||:|||| |
Db        307 KKIQILWPK 315
```

RESULT 2
Q8BQZ3

```
ID   Q8BQZ3      PRELIMINARY;      PRT;   701 AA.
AC   Q8BQZ3;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Adenosine deaminase 3.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Brain;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK046083; BAC32597.1; -.
SQ   SEQUENCE   701 AA;  77260 MW;  E9B076EF9BCF58E2 CRC64;

  Query Match           81.2%;  Score 39;  DB 11;  Length 701;
  Best Local Similarity  75.0%;  Pred. No. 35;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 KLQILWKK 9
              |||::|||
Db        112 KLQLIWKK 119


RESULT 3
Q9JIE5
ID   Q9JIE5      PRELIMINARY;      PRT;   731 AA.
AC   Q9JIE5;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   RNA adenosine deaminase RED2 (Fragment).
GN   ADAR3.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=20314475; PubMed=10854777;
RA   Slavov D., Clark M., Gardiner K.;
RT   "Comparative analysis of the RED1 and RED2 A-to-I RNA editing genes
RT   from mammals, pufferfish and zebrafish.";
RL   Gene 250:41-51(2000).
DR   EMBL; AF232942; AAF78580.1; -.
DR   MGD; MGI:2151118; Adar3.
DR   InterPro; IPR002466; A_deamin.
DR   InterPro; IPR001159; DS_RBD.
DR   Pfam; PF02137; A_deamin; 1.
DR   Pfam; PF00035; dsrm; 2.
DR   SMART; SM00552; ADEAMc; 1.
DR   SMART; SM00358; DSRM; 2.
DR   PROSITE; PS50141; A_DEAMIN_EDITASE; 1.
DR   PROSITE; PS50137; DS_RBD; 2.
FT   NON_TER       1      1
FT   NON_TER     731    731
SQ   SEQUENCE   731 AA;  80857 MW;  B2F8CBE179BF4A89 CRC64;

  Query Match           81.2%;  Score 39;  DB 11;  Length 731;
  Best Local Similarity  75.0%;  Pred. No. 37;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 KLQILWKK 9
              |||::|||
Db        102 KLQLIWKK 109


RESULT 4
Q9JI20
ID   Q9JI20      PRELIMINARY;      PRT;   745 AA.
AC   Q9JI20;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   RNA dependent adenosine deaminase 3.
GN   ADAR3.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J;
RA   Chen M.-H., Kabir K., Yang J.-H.;
RT   "Editing activity of mouse RED2 (ADAR3) in vitro.";
RL   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF270495; AAF76894.1; -.
DR   MGD; MGI:2151118; Adar3.
DR   InterPro; IPR002466; A_deamin.
DR   InterPro; IPR001159; DS_RBD.
DR   Pfam; PF02137; A_deamin; 1.
DR   Pfam; PF00035; dsrm; 2.
DR   SMART; SM00552; ADEAMc; 1.
DR   SMART; SM00358; DSRM; 2.
DR   PROSITE; PS50141; A_DEAMIN_EDITASE; 1.
DR   PROSITE; PS50137; DS_RBD; 2.
SQ   SEQUENCE   745 AA;  82085 MW;  6F208288F741B7D8 CRC64;

  Query Match           81.2%;  Score 39;  DB 11;  Length 745;
  Best Local Similarity  75.0%;  Pred. No. 38;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 KLQILWKK 9
              |||::|||
Db        112 KLQLIWKK 119


RESULT 5
Q8CC51
ID   Q8CC51      PRELIMINARY;      PRT;   745 AA.
AC   Q8CC51;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Adenosine deaminase 3.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Diencephalon;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK033916; BAC28513.1; -.
SQ   SEQUENCE   745 AA;  82187 MW;  A336473C760902F7 CRC64;

  Query Match           81.2%;  Score 39;  DB 11;  Length 745;
  Best Local Similarity  75.0%;  Pred. No. 38;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 KLQILWKK 9
              |||::|||
Db        112 KLQLIWKK 119
```

```
RESULT 6
P97616
ID   P97616      PRELIMINARY;      PRT;   746 AA.
AC   P97616;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Double-stranded RNA specific adenosine deaminase.
GN   RED2.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97112965; PubMed=8943218;
RA   Melcher T., Maas S., Herb A., Sprengel R., Higuchi M., Seeburg P.H.;
RT   "RED2, a brain-specific member of the RNA-specific adenosine deaminase
RT   family.";
RL   J. Biol. Chem. 271:31795-31798(1996).
DR   EMBL; U74586; AAB41862.1; -.
DR   InterPro; IPR002466; A_deamin.
DR   InterPro; IPR001159; DS_RBD.
DR   Pfam; PF02137; A_deamin; 1.
DR   Pfam; PF00035; dsrm; 2.
DR   SMART; SM00552; ADEAMc; 1.
DR   SMART; SM00358; DSRM; 2.
DR   PROSITE; PS50141; A_DEAMIN_EDITASE; 1.
DR   PROSITE; PS50137; DS_RBD; 2.
SQ   SEQUENCE   746 AA;  82227 MW;  6741BA5FDA699077 CRC64;

  Query Match             81.2%;  Score 39;  DB 11;  Length 746;
  Best Local Similarity   75.0%;  Pred. No. 38;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 KLQILWKK 9
              |||::|||
Db        113 KLQLIWKK 120


RESULT 7
Q96BQ3
ID   Q96BQ3      PRELIMINARY;      PRT;   446 AA.
AC   Q96BQ3;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RA   Strausberg R.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR   EMBL; BC015353; AAH15353.1; -.
DR   Genew; HGNC:19015; TRIM43.
DR   InterPro; IPR001870; B302.
DR   InterPro; IPR003877; SPRY_receptor.
DR   InterPro; IPR000315; Znf_Bbox.
DR   InterPro; IPR001841; Znf_ring.
DR   Pfam; PF00622; SPRY; 1.
DR   Pfam; PF00643; zf-B_box; 1.
DR   Pfam; PF00097; zf-C3HC4; 1.
DR   SMART; SM00184; RING; 1.
DR   PROSITE; PS50119; ZF_BBOX; 1.
DR   PROSITE; PS00518; ZF_RING_1; 1.
DR   PROSITE; PS50089; ZF_RING_2; 1.
KW   Hypothetical protein; Metal-binding; Zinc; Zinc-finger.
SQ   SEQUENCE   446 AA;  52265 MW;  805E4AA25617724A CRC64;

  Query Match             79.2%;  Score 38;  DB 4;  Length 446;
  Best Local Similarity   66.7%;  Pred. No. 36;
  Matches    6;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLQILWKK 9
              |:::|||||
Db        142 KQMRILWKK 150


RESULT 8
Q9LSS5
ID   Q9LSS5      PRELIMINARY;      PRT;   564 AA.
AC   Q9LSS5;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   Myosin heavy chain-like.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Columbia;
RA   Kaneko T., Katoh T., Asamizu E., Sato S., Nakamura Y., Kotani H.,
RA   Tabata S.;
RT   "Structural analysis of Arabidopsis thaliana chromosome 5. XI.";
RL   Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB026632; BAA97502.1; -.
SQ   SEQUENCE   564 AA;  63712 MW;  18A9F64BA991F2C9 CRC64;

  Query Match             79.2%;  Score 38;  DB 10;  Length 564;
  Best Local Similarity   66.7%;  Pred. No. 45;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKLQILWKK 9
              ||: :||||
Db        553 KKIGVLWKK 561


RESULT 9
Q9NS39
ID   Q9NS39      PRELIMINARY;      PRT;   739 AA.
AC   Q9NS39;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Double-stranded RNA specific adenosine deaminase.
GN   ADAR3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20294785; PubMed=10836796;
RA   Chen C.X., Cho D.S., Wang Q., Lai F., Carter K.C., Nishikura K.;
RT   "A third member of the RNA-specific adenosine deaminase gene family,
RT   ADAR3, contains both single- and double-stranded RNA binding
RT   domains.";
RL   RNA 6:755-767(2000).
DR   EMBL; AF034837; AAF78094.1; -.
DR   InterPro; IPR002466; A_deamin.
DR   InterPro; IPR001159; DS_RBD.
DR   Pfam; PF02137; A_deamin; 1.
DR   Pfam; PF00035; dsrm; 2.
DR   SMART; SM00552; ADEAMc; 1.
DR   SMART; SM00358; DSRM; 2.
DR   PROSITE; PS50141; A_DEAMIN_EDITASE; 1.
```

```
DR   PROSITE; PS50137; DS_RBD: 2.
SQ   SEQUENCE   739 AA;  80621 MW;  A158A81D35894F79 CRC64;

  Query Match             79.2%;  Score 38;  DB 4;  Length 739;
  Best Local Similarity   75.0%;  Pred. No. 58;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 KLQILWKK 9
              |||::|||
Db        112 KLQLVWKK 119


RESULT 10
Q8I634
ID   Q8I634      PRELIMINARY;      PRT;   900 AA.
AC   Q8I634;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Protein with DNAJ domain (resa-like), putative.
GN   PFL0055C.
OS   Plasmodium falciparum (isolate 3D7).
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=36329;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=3D7;
RX   MEDLINE=22255705; PubMed=12368864;
RA   Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA   Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA   Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA   Chan M.-S., Nene V., Shallom S.J., Suh B., Peterson J., Angiuoli S.,
RA   Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA   Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA   McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA   Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA   Fraser C.M., Barrell B.;
RT   "Genome sequence of the human malaria parasite Plasmodium
RT   falciparum.";
RL   Nature 419:498-511(2002).
DR   EMBL; AE014844; AAN36100.1; -.
SQ   SEQUENCE   900 AA;  107626 MW;  AE90E2273EB5FCAA CRC64;

  Query Match             79.2%;  Score 38;  DB 5;  Length 900;
  Best Local Similarity   55.6%;  Pred. No. 70;
  Matches    5;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLQILWKK 9
              ||::::|||
Db        181 KKIEMIWKK 189


RESULT 11
Q8JNA1
ID   Q8JNA1      PRELIMINARY;      PRT;   148 AA.
AC   Q8JNA1;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Putative transforming protein E6.
GN   E6.
OS   Human papillomavirus type 90.
OC   Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC   Papillomavirus.
OX   NCBI_TaxID=202251;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22079881; PubMed=12085327;
RA   Terai M., Burk R.D.;
RT   "Identification and Characterization of 3 Novel Genital Human
RT   Papillomaviruses by Overlapping Polymerase Chain Reaction: candHPV89,
RT   candHPV90, and candHPV91.";
RL   J. Infect. Dis. 185:1794-1797(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Burk R.D., Terai M.;
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY057438; AAL14204.1; -.
DR   InterPro; IPR000345; CytC_heme_bind.
DR   InterPro; IPR001334; E6.
DR   InterPro; IPR000566; Lipocln_cytFABP.
DR   Pfam; PF00518; E6; 1.
DR   PROSITE; PS00190; CYTOCHROME_C; 1.
DR   PROSITE; PS00213; LIPOCALIN; 1.
SQ   SEQUENCE   148 AA;  17173 MW;  374D44C7D05C95DA CRC64;

  Query Match             77.1%;  Score 37;  DB 12;  Length 148;
  Best Local Similarity   55.6%;  Pred. No. 20;
  Matches    5;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLQILWKK 9
              |:||::|:|
Db         47 KELQVVWRK 55


RESULT 12
Q9BUW4
ID   Q9BUW4      PRELIMINARY;      PRT;   208 AA.
AC   Q9BUW4;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Uterus;
RA   Strausberg R.;
RL   Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR   EMBL; BC001862; AAH01862.1; -.
DR   Genew; HGNC:19021; TRIM48.
DR   InterPro; IPR000315; Znf_Bbox.
DR   InterPro; IPR001841; Znf_ring.
DR   Pfam; PF00643; zf-B_box; 1.
DR   Pfam; PF00097; zf-C3HC4; 1.
DR   SMART; SM00336; BBOX; 1.
DR   SMART; SM00184; RING; 1.
DR   PROSITE; PS50119; ZF_BBOX; 1.
DR   PROSITE; PS50089; ZF_RING_2; 1.
KW   Hypothetical protein; Metal-binding; Zinc; Zinc-finger.
SQ   SEQUENCE   208 AA;  24497 MW;  93DD7840775F6E57 CRC64;

  Query Match             75.0%;  Score 36;  DB 4;  Length 208;
  Best Local Similarity   66.7%;  Pred. No. 42;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKLQILWKK 9
              ||:| ||:|
Db        142 KKMQSLWEK 150


RESULT 13
Q8XMB9
ID   Q8XMB9      PRELIMINARY;      PRT;   279 AA.
AC   Q8XMB9;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Probable transcriptional regulator.
GN   CPE0770.
OS   Clostridium perfringens.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
```

```
OC   Clostridium.
OX   NCBI_TaxID=1502;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=13 / Type A;
RX   PubMed=11792842;
RA   Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
RA   Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
RT   "Complete genome sequence of Clostridium perfringens, an anaerobic
RT   flesh-eater.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
DR   EMBL; AP003188; BAB80476.1; -.
DR   InterPro; IPR000005; HTHAraC.
DR   Pfam; PF00165; HTH_AraC; 2.
DR   PRINTS; PR00032; HTHARAC.
DR   SMART; SM00342; HTH_ARAC; 1.
DR   PROSITE; PS00041; HTH_ARAC_FAMILY_1; 1.
DR   PROSITE; PS01124; HTH_ARAC_FAMILY_2; 1.
KW   Complete proteome.
SQ   SEQUENCE   279 AA;  32341 MW;  F511D10089D409E4 CRC64;

  Query Match             75.0%;  Score 36;  DB 16;  Length 279;
  Best Local Similarity   85.7%;  Pred. No. 56;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LQILWKK 9
              :|||||| 
Db          1 MQILWKK 7


RESULT 14
Q8IWZ4
ID   Q8IWZ4      PRELIMINARY;      PRT;   375 AA.
AC   Q8IWZ4;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Tripartite motif protein 48.
GN   TRIM48.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Meroni G.;
RT   "Novel tripartite motif family members.";
RL   Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF521869; AAO14946.1; -.
SQ   SEQUENCE   375 AA;  43408 MW;  03C69CB0453C2FD8 CRC64;

  Query Match             75.0%;  Score 36;  DB 4;  Length 375;
  Best Local Similarity   66.7%;  Pred. No. 74;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKLQILWKK 9
              ||:| ||:|
Db        142 KKMQSLWEK 150


RESULT 15
Q8K5B1
ID   Q8K5B1      PRELIMINARY;      PRT;   716 AA.
AC   Q8K5B1;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Gp130-like monocyte receptor.
GN   GLMR.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21988187; PubMed=11877449;
RA   Ghilardi N., Li J., Hongo J.A., Yi S., Gurney A., De Sauvage F.J.;
RT   "A Novel Type I Cytokine Receptor Is Expressed on Monocytes, Signals
RT   Proliferation, and Activates STAT-3 and STAT-5.";
RL   J. Biol. Chem. 277:16831-16836(2002).
DR   EMBL; AF486621; AAM27959.1; -.
DR   MGD; MGI:2180511; Glmr.
DR   InterPro; IPR002996; CR1A.
DR   InterPro; IPR003961; FN_III.
DR   Pfam; PF00041; fn3; 1.
DR   SMART; SM00060; FN3; 3.
KW   Receptor.
SQ   SEQUENCE   716 AA;  80641 MW;  39E6B0B253F7C7E5 CRC64;

  Query Match             75.0%;  Score 36;  DB 11;  Length 716;
  Best Local Similarity   55.6%;  Pred. No. 1.4e+02;
  Matches    5;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLQILWKK 9
              :|:::|||| 
Db        231 RKVRLLWKK 239


Search completed: October 14, 2003, 14:36:58
Job time : 32.8125 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 14:33:19 ; Search time 38.625 Seconds
(without alignments)
36.985 Million cell updates/sec

Title: US-09-881-490-164
Perfect score: 48
Sequence: 1 KKWLIQLKK 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	48	100.0	9	17	AAW04030	Antifungal peptide
2	48	100.0	9	18	AAW44559	Anti-fungal peptid
3	48	100.0	9	18	AAW44561	Anti-fungal peptid
4	48	100.0	9	18	AAW44563	Anti-fungal peptid
5	48	100.0	9	18	AAW44572	Anti-fungal peptid
6	48	100.0	9	18	AAW43753	Bactericidal/perme
7	48	100.0	9	18	AAW43749	Bactericidal/perme
8	48	100.0	9	20	AAY00536	Antifungal peptide
9	48	100.0	9	20	AAY00538	Antifungal peptide
10	48	100.0	9	20	AAY00540	Antifungal peptide
11	48	100.0	9	20	AAY00549	Antifungal peptide
12	48	100.0	9	21	AAY88384	Antifungal peptide
13	48	100.0	9	22	AAB65460	Anti-fungal peptid
14	48	100.0	9	22	AAB65462	Anti-fungal peptid
15	48	100.0	9	22	AAB65464	Anti-fungal peptid
16	48	100.0	9	22	AAB65473	Anti-fungal peptid
17	48	100.0	10	17	AAW04039	Antifungal peptide
18	48	100.0	10	17	AAW04040	Antifungal peptide
19	48	100.0	10	18	AAW44578	Anti-fungal peptid
20	48	100.0	10	18	AAW44579	Anti-fungal peptid
21	48	100.0	10	18	AAW43761	Bactericidal/perme
22	48	100.0	10	18	AAW43760	Bactericidal/perme
23	48	100.0	10	20	AAY00555	Antifungal peptide
24	48	100.0	10	20	AAY00556	Antifungal peptide
25	48	100.0	10	22	AAB65479	Anti-fungal peptid
26	48	100.0	10	22	AAB65480	Anti-fungal peptid
27	48	100.0	11	17	AAW04041	Antifungal peptide
28	48	100.0	11	18	AAW44580	Anti-fungal peptid
29	48	100.0	11	18	AAW43762	Bactericidal/perme
30	48	100.0	11	20	AAY00557	Antifungal peptide
31	48	100.0	11	22	AAB65481	Anti-fungal peptid
32	43	89.6	8	17	AAW04027	Antifungal peptide
33	43	89.6	8	17	AAW04029	Antifungal peptide
34	43	89.6	8	18	AAW44556	Anti-fungal peptid
35	43	89.6	8	18	AAW44558	Anti-fungal peptid
36	43	89.6	8	18	AAW43748	Bactericidal/perme
37	43	89.6	8	20	AAY00535	Antifungal peptide
38	43	89.6	8	20	AAY00533	Antifungal peptide
39	43	89.6	8	22	AAB65457	Anti-fungal peptid
40	43	89.6	8	22	AAB65459	Anti-fungal peptid
41	43	89.6	9	17	AAW04031	Antifungal peptide
42	43	89.6	9	17	AAW04036	Antifungal peptide
43	43	89.6	9	17	AAW04037	Antifungal peptide
44	43	89.6	9	18	AAW44564	Anti-fungal peptid
45	43	89.6	9	18	AAW44565	Anti-fungal peptid

ALIGNMENTS

RESULT 1
AAW04030
ID AAW04030 standard; peptide; 9 AA.
XX
AC AAW04030;
XX
DT 04-NOV-1996 (first entry)
XX
DE Antifungal peptide XMP.327/XMP.331/XMP.333/XMP.342.
XX
KW Antifungal peptide; inhibitor; Domain III; polymorphonuclear leukocyte;
KW bactericidal/permeability-increasing protein; BPI; mammalian; PMN; fungi;
KW neutrophil; replication inhibitor; fungal infection; Aspergillus;
KW Cryptococcus; Candida; C.albicans; C.galabrat; C.krusei; C.lusitaniae;
KW C.parapsilosis; C.tropicalis; therapy.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "optionally acetylated, optionally D-form residue"
FT Misc-difference 2
FT /note= "optionally D-form residue"
FT Misc-difference 8
FT /note= "optionally D-form residue"
FT Modified-site 9
FT /note= "amidated, optionally D-form residue"
XX
PN WO9608509-A1.
XX
PD 21-MAR-1996.
XX

PF 20-JUL-1995; 95WO-US09262.
XX
PR 13-JAN-1995; 95US-0372105.
PR 15-SEP-1994; 94US-0306473.
XX
PA (XOMA) XOMA CORP.
XX
PI Fadem MB, Lim E, Little RG;
XX
DR WPI; 1996-179900/18.
XX
PT Antifungal peptide(s) derived from Domain III of BPI protein - used
PT in vitro for killing or inhibiting replication of fungi, esp.
PT Candida species
XX
PS Claim 5; Page 154; 199pp; English.
XX
CC AAW04000-W04160 represent antifungal peptides. These sequences are
CC based on (or derived from) Domain III of the
CC bactericidal/permeability-increasing protein (BPI). BPI is a protein
CC that can be isolated from the granules of mammalian polymorphonuclear
CC leukocytes (PMNs or neutrophils). These antifungal peptides can be used
CC for killing, or inhibiting replication of, fungi in vitro. These
CC sequences can also be used for treatment of a fungal infection involving
CC fungi from the species Candida, Aspergillus and Cryptococcus. The
CC sequences are especially useful for treating C.albicans, C.galabrat,
CC C.krusei, C.lusitaniae, C.parapsilosis and C.tropicalis infections.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 48; DB 17; Length 9;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKWLIQLKK 9
||||||||| |
Db 1 KKWLIQLKK 9

RESULT 2
AAW44559
ID AAW44559 standard; peptide; 9 AA.
XX
AC AAW44559;
XX
DT 27-APR-1998 (first entry)
XX
DE Anti-fungal peptide #160 based on BPI protein (residues 142-169).
XX
KW Anti-fungal peptide; bactericidal-permeability-increasing protein; BPI;
KW polymorphonuclear leukocyte; fungicide.
XX
OS Synthetic.
OS Mammalia.
XX
FH Key Location/Qualifiers
FT Modified-site 9
FT /note= "C-terminal amide"
XX
PN WO9704008-A1.
XX
PD 06-FEB-1997.
XX
PF 21-MAR-1996; 96WO-US03845.
XX
PR 20-JUL-1995; 95US-0504841.
XX
PA (XOMA) XOMA CORP.
XX
PI Fadem MB, Lim E, Little RG;
XX
DR WPI; 1997-132578/12.
XX
PT Anti-fungal peptide(s) derived from or based on domain III of
PT bactericidal-permeability-increasing protein - are used in vitro or
PT in vivo as a fungicides
XX
PS Claim 1; Page 191; 230pp; English.
XX
CC This is a specifically claimed anti-fungal peptide which is based on
CC domain III (amino acids 142-160) of bactericidal-permeability-increasing
CC protein (BPI), isolated from the granules of mammalian polymorphonuclear
CC leukocytes. It is used in compositions with diluents, carriers or
CC adjuvants to treat fungal infections in patients. It may also be used in
CC vitro to kill or inhibit the replication of fungi, such as in
CC decontaminating fluids and sterilising medical and implant devices.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 48; DB 18; Length 9;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKWLIQLKK 9
||||||||| |
Db 1 KKWLIQLKK 9

RESULT 3
AAW44561
ID AAW44561 standard; peptide; 9 AA.
XX
AC AAW44561;
XX
DT 27-APR-1998 (first entry)
XX
DE Anti-fungal peptide #162 based on BPI protein (residues 142-169).
XX
KW Anti-fungal peptide; bactericidal-permeability-increasing protein; BPI;
KW polymorphonuclear leukocyte; fungicide.
XX
OS Synthetic.
OS Mammalia.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "N-terminal acetyl"
FT Modified-site 9
FT /note= "C-terminal amide"
XX
PN WO9704008-A1.
XX
PD 06-FEB-1997.
XX
PF 21-MAR-1996; 96WO-US03845.
XX
PR 20-JUL-1995; 95US-0504841.
XX
PA (XOMA) XOMA CORP.
XX
PI Fadem MB, Lim E, Little RG;
XX
DR WPI; 1997-132578/12.
XX
PT Anti-fungal peptide(s) derived from or based on domain III of
PT bactericidal-permeability-increasing protein - are used in vitro or
PT in vivo as a fungicides
XX
PS Claim 1; Page 192; 230pp; English.
XX
CC This is a specifically claimed anti-fungal peptide which is based on
CC domain III (amino acids 142-160) of bactericidal-permeability-increasing
CC protein (BPI), isolated from the granules of mammalian polymorphonuclear
CC leukocytes. It is used in compositions with diluents, carriers or
CC adjuvants to treat fungal infections in patients. It may also be used in
CC vitro to kill or inhibit the replication of fungi, such as in

```
CC   decontaminating fluids and sterilising medical and implant devices.
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 48;  DB 18;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 9.3e-05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


RESULT 4
AAW44563
ID   AAW44563 standard; peptide; 9 AA.
XX
AC   AAW44563;
XX
DT   27-APR-1998  (first entry)
XX
DE   Anti-fungal peptide #164 based on BPI protein (residues 142-169).
XX
KW   Anti-fungal peptide; bactericidal-permeability-increasing protein; BPI;
KW   polymorphonuclear leukocyte; fungicide.
XX
OS   Synthetic.
OS   Mammalia.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1
FT                   /note= "D-form residue"
FT   Modified-site   9
FT                   /note= "C-terminal amide"
XX
PN   WO9704008-A1.
XX
PD   06-FEB-1997.
XX
PF   21-MAR-1996;   96WO-US03845.
XX
PR   20-JUL-1995;   95US-0504841.
XX
PA   (XOMA ) XOMA CORP.
XX
PI   Fadem MB,  Lim E,  Little RG;
XX
DR   WPI; 1997-132578/12.
XX
PT   Anti-fungal peptide(s) derived from or based on domain III of
PT   bactericidal-permeability-increasing protein - are used in vitro or
PT   in vivo as a fungicides
XX
PS   Claim 1; Page 192; 230pp; English.
XX
CC   This is a specifically claimed anti-fungal peptide which is based on
CC   domain III (amino acids 142-160) of bactericidal-permeability-increasing
CC   protein (BPI), isolated from the granules of mammalian polymorphonuclear
CC   leukocytes. It is used in compositions with diluents, carriers or
CC   adjuvants to treat fungal infections in patients. It may also be used in
CC   vitro to kill or inhibit the replication of fungi, such as in
CC   decontaminating fluids and sterilising medical and implant devices.
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 48;  DB 18;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 9.3e-05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


RESULT 5
AAW44572
ID   AAW44572 standard; peptide; 9 AA.
XX
AC   AAW44572;
XX
DT   27-APR-1998  (first entry)
XX
DE   Anti-fungal peptide #173 based on BPI protein (residues 142-169).
XX
KW   Anti-fungal peptide; bactericidal-permeability-increasing protein; BPI;
KW   polymorphonuclear leukocyte; fungicide.
XX
OS   Synthetic.
OS   Mammalia.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1
FT                   /note= "D-form residue"
FT   Misc-difference 2
FT                   /note= "D-form residue"
FT   Misc-difference 8
FT                   /note= "D-form residue"
FT   Misc-difference 9
FT                   /note= "D-form residue, C-terminal amide"
XX
PN   WO9704008-A1.
XX
PD   06-FEB-1997.
XX
PF   21-MAR-1996;   96WO-US03845.
XX
PR   20-JUL-1995;   95US-0504841.
XX
PA   (XOMA ) XOMA CORP.
XX
PI   Fadem MB,  Lim E,  Little RG;
XX
DR   WPI; 1997-132578/12.
XX
PT   Anti-fungal peptide(s) derived from or based on domain III of
PT   bactericidal-permeability-increasing protein - are used in vitro or
PT   in vivo as a fungicides
XX
PS   Claim 1; Page 196; 230pp; English.
XX
CC   This is a specifically claimed anti-fungal peptide which is based on
CC   domain III (amino acids 142-160) of bactericidal-permeability-increasing
CC   protein (BPI), isolated from the granules of mammalian polymorphonuclear
CC   leukocytes. It is used in compositions with diluents, carriers or
CC   adjuvants to treat fungal infections in patients. It may also be used in
CC   vitro to kill or inhibit the replication of fungi, such as in
CC   decontaminating fluids and sterilising medical and implant devices.
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 48;  DB 18;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 9.3e-05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


RESULT 6
AAW43753
ID   AAW43753 standard; peptide; 9 AA.
XX
AC   AAW43753;
XX
DT   20-APR-1998  (first entry)
```

XX
DE Bactericidal/permeability increasing peptide XMP.331.
XX
KW Bactericidal/permeability increasing peptide; BPI; fusion protein;
KW bacterial infection; fungal infection; endotoxin; heparin;
KW angiogenesis; fungicidal; recombinant DNA; vector.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "Acetylated"
FT Modified-site 9
FT /note= "Amidated"
XX
PN WO9735009-A1.
XX
PD 25-SEP-1997.
XX
PF 18-MAR-1997; 97WO-US05287.
XX
PR 22-MAR-1996; 96US-0621803.
XX
PA (XOMA) XOMA CORP.
XX
PI Better MD;
XX
DR WPI; 1997-480215/44.
XX
PT Recombinant production of bactericidal/permeability increasing
PT protein - by expression as a fusion protein in microbial host cells,
PT then cleaving the BPI peptide from the carrier
XX
PS Claim 10; Page 127; 186pp; English.
XX
CC A new recombinant DNA vector construct has been developed which encodes
CC a fusion protein and is suitable for introduction into a bacterial host.
CC The vector comprises: (a) DNA encoding at least one cationic
CC bactericidal/permeability increasing peptide (BPI), (b) DNA encoding a
CC carrier protein, and (c) DNA encoding an amino acid (aa) cleavage site
CC located between (a) and (b). The present sequence represents a
CC specifically claimed BPI peptide. The peptides have many uses including
CC the treatment of bacterial and fungal infections. BPI peptides also
CC bind to endotoxins and heparin, neutralising their effects. The
CC peptides have further been shown to inhibit angiogenesis (partly due to
CC heparin-binding activity). The fusion proteins have been found to be
CC expressed in large amounts without significant proteolysis, and in some
CC cases are actually secreted from the host cells. This allows the
CC indirect production of anti-microbial BPI peptides in microbial hosts.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 48; DB 18; Length 9;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKWLIQLKK 9
||||||||| |
Db 1 KKWLIQLKK 9

RESULT 7
AAW43749
ID AAW43749 standard; peptide; 9 AA.
XX
AC AAW43749;
XX
DT 20-APR-1998 (first entry)
XX
DE Bactericidal/permeability increasing peptide XMP.327.
XX
KW Bactericidal/permeability increasing peptide; BPI; fusion protein;
KW bacterial infection; fungal infection; endotoxin; heparin;
KW angiogenesis; fungicidal; recombinant DNA; vector.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 9
FT /note= "Amidated"
XX
PN WO9735009-A1.
XX
PD 25-SEP-1997.
XX
PF 18-MAR-1997; 97WO-US05287.
XX
PR 22-MAR-1996; 96US-0621803.
XX
PA (XOMA) XOMA CORP.
XX
PI Better MD;
XX
DR WPI; 1997-480215/44.
XX
PT Recombinant production of bactericidal/permeability increasing
PT protein - by expression as a fusion protein in microbial host cells,
PT then cleaving the BPI peptide from the carrier
XX
PS Claim 10; Page 125; 186pp; English.
XX
CC A new recombinant DNA vector construct has been developed which encodes
CC a fusion protein and is suitable for introduction into a bacterial host.
CC The vector comprises: (a) DNA encoding at least one cationic
CC bactericidal/permeability increasing peptide (BPI), (b) DNA encoding a
CC carrier protein, and (c) DNA encoding an amino acid (aa) cleavage site
CC located between (a) and (b). The present sequence represents a
CC specifically claimed BPI peptide. The peptides have many uses including
CC the treatment of bacterial and fungal infections. BPI peptides also
CC bind to endotoxins and heparin, neutralising their effects. The
CC peptides have further been shown to inhibit angiogenesis (partly due to
CC heparin-binding activity). The fusion proteins have been found to be
CC expressed in large amounts without significant proteolysis, and in some
CC cases are actually secreted from the host cells. This allows the
CC indirect production of anti-microbial BPI peptides in microbial hosts.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 48; DB 18; Length 9;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKWLIQLKK 9
||||||||| |
Db 1 KKWLIQLKK 9

RESULT 8
AAY00536
ID AAY00536 standard; Peptide; 9 AA.
XX
AC AAY00536;
XX
DT 07-MAY-1999 (first entry)
XX
DE Antifungal peptide XMP.327.
XX
KW Antifungal; BPI; bactericidal/permeability increasing protein;
KW Candida infection.
XX
OS Synthetic.
XX
PN US5858974-A.
XX

PD 12-JAN-1999.
XX
PF 21-MAR-1996; 96US-0621259.
XX
PR 21-MAR-1996; 96US-0621259.
PR 20-JUL-1995; 95US-0504841.
XX
PA (XOMA) XOMA CORP.
XX
PI Fadem MB, Lim E, Little RG;
XX
DR WPI; 1999-119956/10.
XX
PT Antifungal peptides - comprising part of bactericidal or
PT permeability-increasing protein sequence or related sequence
XX
PS Disclosure; Columns 165-166; 132pp; English.
XX
CC New peptides are provided which are based on Domain III (amino acids
CC 142-169) of human bactericidal/permeability-increasing protein (BPI).
CC The peptides all have a C-terminal amide. More particularly, the Claims
CC relate to: (1) a peptide that has an amino acid sequence of human BPI
CC from position 148 to position 161 (KSKVGWLIQLFHKK) and variants of the
CC sequence having antifungal activity; and (2) an antifungal peptide
CC having 6-14 amino acids comprising (a) a core sequence selected from
CC LIQL, IQLF, WLIQL, LIQLF and WLIQLF and (b) one or more cationic amino
CC acids selected from K, R, H, Orn (ornithine) and Dab (diaminobutyric
CC acid) at the N and/or C terminus of the core sequence. The new peptides
CC are used for killing or inhibiting replication of fungi in vitro; and
CC for treating fungal infections in vivo, in particular infections of
CC Candida, Aspergillus or Cryptococcus spp., especially C. albicans, C.
CC krusei, C. lusitaniae, C. parapsilosis or C. tropicalis. The peptide
CC can be administered topically, intravenously, orally or as an aerosol,
CC optionally together with a non-peptide antifungal agent.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 48; DB 20; Length 9;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKWLIQLKK 9
|||||||||
Db 1 KKWLIQLKK 9

RESULT 9
AAY00538
ID AAY00538 standard; Peptide; 9 AA.
XX
AC AAY00538;
XX
DT 07-MAY-1999 (first entry)
XX
DE Antifungal peptide XMP.331.
XX
KW Antifungal; BPI; bactericidal/permeability increasing protein;
KW Candida infection.
XX
OS Synthetic.
XX
PN US5858974-A.
XX
PD 12-JAN-1999.
XX
PF 21-MAR-1996; 96US-0621259.
XX
PR 21-MAR-1996; 96US-0621259.
PR 20-JUL-1995; 95US-0504841.
XX
PA (XOMA) XOMA CORP.
XX
PI Fadem MB, Lim E, Little RG;
XX
DR WPI; 1999-119956/10.
XX
PT Antifungal peptides - comprising part of bactericidal or
PT permeability-increasing protein sequence or related sequence
XX
PS Disclosure; Columns 165-166; 132pp; English.
XX
CC New peptides are provided which are based on Domain III (amino acids
CC 142-169) of human bactericidal/permeability-increasing protein (BPI).
CC The peptides all have a C-terminal amide. More particularly, the Claims
CC relate to: (1) a peptide that has an amino acid sequence of human BPI
CC from position 148 to position 161 (KSKVGWLIQLFHKK) and variants of the
CC sequence having antifungal activity; and (2) an antifungal peptide
CC having 6-14 amino acids comprising (a) a core sequence selected from
CC LIQL, IQLF, WLIQL, LIQLF and WLIQLF and (b) one or more cationic amino
CC acids selected from K, R, H, Orn (ornithine) and Dab (diaminobutyric
CC acid) at the N and/or C terminus of the core sequence. The new peptides
CC are used for killing or inhibiting replication of fungi in vitro; and
CC for treating fungal infections in vivo, in particular infections of
CC Candida, Aspergillus or Cryptococcus spp., especially C. albicans, C.
CC krusei, C. lusitaniae, C. parapsilosis or C. tropicalis. The peptide
CC can be administered topically, intravenously, orally or as an aerosol,
CC optionally together with a non-peptide antifungal agent.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 48; DB 20; Length 9;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKWLIQLKK 9
|||||||||
Db 1 KKWLIQLKK 9

RESULT 10
AAY00540
ID AAY00540 standard; Peptide; 9 AA.
XX
AC AAY00540;
XX
DT 07-MAY-1999 (first entry)
XX
DE Antifungal peptide XMP.333.
XX
KW Antifungal; BPI; bactericidal/permeability increasing protein;
KW Candida infection.
XX
OS Synthetic.
XX
PN US5858974-A.
XX
PD 12-JAN-1999.
XX
PF 21-MAR-1996; 96US-0621259.
XX
PR 21-MAR-1996; 96US-0621259.
PR 20-JUL-1995; 95US-0504841.
XX
PA (XOMA) XOMA CORP.
XX
PI Fadem MB, Lim E, Little RG;
XX
DR WPI; 1999-119956/10.
XX
PT Antifungal peptides - comprising part of bactericidal or
PT permeability-increasing protein sequence or related sequence
XX
PS Disclosure; Columns 167-168; 132pp; English.
XX
CC New peptides are provided which are based on Domain III (amino acids
CC 142-169) of human bactericidal/permeability-increasing protein (BPI).

```
CC   The peptides all have a C-terminal amide. More particularly, the Claims
CC   relate to: (1) a peptide that has an amino acid sequence of human BPI
CC   from position 148 to position 161 (KSKVGWLIQLFHKK) and variants of the
CC   sequence having antifungal activity; and (2) an antifungal peptide
CC   having 6-14 amino acids comprising (a) a core sequence selected from
CC   LIQL, IQLF, WLIQL, LIQLF and WLIQLF and (b) one or more cationic amino
CC   acids selected from K, R, H, Orn (ornithine) and Dab (diaminobutyric
CC   acid) at the N and/or C terminus of the core sequence. The new peptides
CC   are used for killing or inhibiting replication of fungi in vitro; and
CC   for treating fungal infections in vivo, in particular infections of
CC   Candida, Aspergillus or Cryptococcus spp., especially C. albicans, C.
CC   krusei, C. lusitaniae, C. parapsilosis or C. tropicalis. The peptide
CC   can be administered topically, intravenously, orally or as an aerosol,
CC   optionally together with a non-peptide antifungal agent.
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 48;  DB 20;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 9.3e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


RESULT 11
AAY00549
ID   AAY00549 standard; Peptide; 9 AA.
XX
AC   AAY00549;
XX
DT   07-MAY-1999  (first entry)
XX
DE   Antifungal peptide XMP.342.
XX
KW   Antifungal; BPI; bactericidal/permeability increasing protein;
KW   Candida infection.
XX
OS   Synthetic.
XX
PN   US5858974-A.
XX
PD   12-JAN-1999.
XX
PF   21-MAR-1996;   96US-0621259.
XX
PR   21-MAR-1996;   96US-0621259.
PR   20-JUL-1995;   95US-0504841.
XX
PA   (XOMA ) XOMA CORP.
XX
PI   Fadem MB,  Lim E,  Little RG;
XX
DR   WPI; 1999-119956/10.
XX
PT   Antifungal peptides - comprising part of bactericidal or
PT   permeability-increasing protein sequence or related sequence
XX
PS   Disclosure; Columns 173-174; 132pp; English.
XX
CC   New peptides are provided which are based on Domain III (amino acids
CC   142-169) of human bactericidal/permeability-increasing protein (BPI).
CC   The peptides all have a C-terminal amide. More particularly, the Claims
CC   relate to: (1) a peptide that has an amino acid sequence of human BPI
CC   from position 148 to position 161 (KSKVGWLIQLFHKK) and variants of the
CC   sequence having antifungal activity; and (2) an antifungal peptide
CC   having 6-14 amino acids comprising (a) a core sequence selected from
CC   LIQL, IQLF, WLIQL, LIQLF and WLIQLF and (b) one or more cationic amino
CC   acids selected from K, R, H, Orn (ornithine) and Dab (diaminobutyric
CC   acid) at the N and/or C terminus of the core sequence. The new peptides
CC   are used for killing or inhibiting replication of fungi in vitro; and
CC   for treating fungal infections in vivo, in particular infections of
CC   Candida, Aspergillus or Cryptococcus spp., especially C. albicans, C.
CC   krusei, C. lusitaniae, C. parapsilosis or C. tropicalis. The peptide
CC   can be administered topically, intravenously, orally or as an aerosol,
CC   optionally together with a non-peptide antifungal agent.
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 48;  DB 20;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 9.3e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


RESULT 12
AAY88384
ID   AAY88384 standard; peptide; 9 AA.
XX
AC   AAY88384;
XX
DT   24-JUL-2000  (first entry)
XX
DE   Antifungal peptide XMP.342.
XX
KW   Antifungal peptide; XMP.342; BPI; microorganism; defence; antibacterial;
KW   bactericidal/permeability increasing protein; antimicrobial; XMP. 445;
KW   fungal infection; Candida; Aspergillus; Fusarium; growth inhibitor;
KW   decontaminate; sterilize; medical equipment; prosthetic joint.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1
FT                   /note= "D-form residue"
FT   Misc-difference 2
FT                   /note= "D-form residue"
FT   Misc-difference 8
FT                   /note= "D-form residue"
FT   Modified-site   9
FT                   /note= "D-form residue, C-terminal amide"
XX
PN   WO200018798-A1.
XX
PD   06-APR-2000.
XX
PF   24-SEP-1999;   99WO-US22277.
XX
PR   25-SEP-1998;   98US-0101958.
PR   25-NOV-1998;   98US-0109896.
XX
PA   (XOMA ) XOMA TECHNOLOGY LTD.
XX
PI   Little RG;
XX
DR   WPI; 2000-303442/26.
XX
PT   A new antifungal peptide, designated XMP.445, used to treat fungal
PT   infections is based on Domain III of
PT   bactericidal/permeability-increasing protein  -
XX
PS   Example 2; Page 28; 31pp; English.
XX
CC   This sequence represents an antifungal peptide designated XMP.342, which
CC   is based on amino acids 153-157 of bactericidal/permeability increasing
CC   (BPI) protein (see AAY88382). BPI is a protein isolated from the
CC   granules of mammalian polymorphonuclear leukocytes (PMNs or neutrophils),
CC   which are essential in the defence against invading microorganisms. BPI
CC   protein products have antibacterial and antimicrobial activity. The
CC   invention relates to antibacterial peptides, specifically XMP.445 (see
CC   AAY88381). The peptide is used to treat fungal infections, particularly
CC   infections of the species Candida, Aspergillus or Fusarium. The peptide
```

CC can also be used to treat a bacterial infection, or to kill or inhibit
CC the growth of bacteria in vitro. The peptide may be useful to treat
CC animals or plants in vivo, and as an agent to decontaminate fluids and
CC surfaces and to sterilize surgical or other medical equipment and
CC implantable devices including prosthetic joints.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 48; DB 21; Length 9;
Best Local Similarity 100.0%; Pred. No. 9.3e-05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKWLIQLKK 9
|||||||||
Db 1 KKWLIQLKK 9

RESULT 13
AAB65460
ID AAB65460 standard; Peptide; 9 AA.
XX
AC AAB65460;
XX
DT 27-MAR-2001 (first entry)
XX
DE Anti-fungal peptide XMP.327.
XX
KW Human; BPI; antifungal; polymorphonuclear leukocyte; neutrophil;
KW bactericidal/permeability-increasing protein; bactericidal;
KW fungal infection.
XX
OS Homo sapiens.
XX
PN US6156730-A.
XX
PD 05-DEC-2000.
XX
PF 08-JAN-1999; 99US-0227659.
XX
PR 21-MAR-1996; 96US-0621259.
PR 12-MAR-1993; 93US-0030644.
PR 15-JUL-1993; 93US-0093202.
PR 14-JAN-1994; 94US-0183222.
PR 11-MAR-1994; 94US-0209762.
PR 11-JUL-1994; 94US-0273540.
PR 15-SEP-1994; 94US-0306473.
PR 13-JAN-1995; 95US-0372105.
PR 20-JUL-1995; 95US-0504841.
XX
PA (XOMA) XOMA CORP.
XX
PI Lim E, Fadem MB, Little RG;
XX
DR WPI; 2001-090160/10.
XX
PT Novel anti-fungal peptides derived from domain III of
PT bactericidal/permeability-increasing protein useful for killing or
PT inhibiting replication of fungi and for treating fungal infections -
XX
PS Example 2; Columns 169-170; 134pp; English.
XX
CC The present invention relates to antifungal peptides (see
CC AAB65301-B65550) derived from Domain III (amino acids 142-169) of
CC bactericidal/permeability-increasing protein (BPI). The present sequence
CC is one such antifungal peptide. BPI is a protein isolated from the
CC granules of mammalian polymorphonuclear leukocytes (PMNs or
CC neutrophils). BPI has potent bactericidal activity against a broad range
CC of gram-negative bacteria. The peptides of the present invention are
CC useful for killing or inhibiting replication of fungi, and treating
CC infections caused by fungus belonging to Candida, Aspergillus,
CC Cryptococcus species such as C.albicans, C.glabrata, C.krusei,
CC C.lusitaniae, C.parapsilosis and C.tropicalis.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 48; DB 22; Length 9;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKWLIQLKK 9
|||||||||
Db 1 KKWLIQLKK 9

RESULT 14
AAB65462
ID AAB65462 standard; Peptide; 9 AA.
XX
AC AAB65462;
XX
DT 27-MAR-2001 (first entry)
XX
DE Anti-fungal peptide XMP.331.
XX
KW Human; BPI; antifungal; polymorphonuclear leukocyte; neutrophil;
KW bactericidal/permeability-increasing protein; bactericidal;
KW fungal infection.
XX
OS Homo sapiens.
XX
PN US6156730-A.
XX
PD 05-DEC-2000.
XX
PF 08-JAN-1999; 99US-0227659.
XX
PR 21-MAR-1996; 96US-0621259.
PR 12-MAR-1993; 93US-0030644.
PR 15-JUL-1993; 93US-0093202.
PR 14-JAN-1994; 94US-0183222.
PR 11-MAR-1994; 94US-0209762.
PR 11-JUL-1994; 94US-0273540.
PR 15-SEP-1994; 94US-0306473.
PR 13-JAN-1995; 95US-0372105.
PR 20-JUL-1995; 95US-0504841.
XX
PA (XOMA) XOMA CORP.
XX
PI Lim E, Fadem MB, Little RG;
XX
DR WPI; 2001-090160/10.
XX
PT Novel anti-fungal peptides derived from domain III of
PT bactericidal/permeability-increasing protein useful for killing or
PT inhibiting replication of fungi and for treating fungal infections -
XX
PS Example 2; Columns 171-172; 134pp; English.
XX
CC The present invention relates to antifungal peptides (see
CC AAB65301-B65550) derived from Domain III (amino acids 142-169) of
CC bactericidal/permeability-increasing protein (BPI). The present sequence
CC is one such antifungal peptide. BPI is a protein isolated from the
CC granules of mammalian polymorphonuclear leukocytes (PMNs or
CC neutrophils). BPI has potent bactericidal activity against a broad range
CC of gram-negative bacteria. The peptides of the present invention are
CC useful for killing or inhibiting replication of fungi, and treating
CC infections caused by fungus belonging to Candida, Aspergillus,
CC Cryptococcus species such as C.albicans, C.glabrata, C.krusei,
CC C.lusitaniae, C.parapsilosis and C.tropicalis.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 48; DB 22; Length 9;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


RESULT 15
AAB65464
ID   AAB65464 standard; Peptide; 9 AA.
XX
AC   AAB65464;
XX
DT   27-MAR-2001  (first entry)
XX
DE   Anti-fungal peptide XMP.333.
XX
KW   Human; BPI; antifungal; polymorphonuclear leukocyte; neutrophil;
KW   bactericidal/permeability-increasing protein; bactericidal;
KW   fungal infection.
XX
OS   Homo sapiens.
XX
PN   US6156730-A.
XX
PD   05-DEC-2000.
XX
PF   08-JAN-1999;   99US-0227659.
XX
PR   21-MAR-1996;   96US-0621259.
PR   12-MAR-1993;   93US-0030644.
PR   15-JUL-1993;   93US-0093202.
PR   14-JAN-1994;   94US-0183222.
PR   11-MAR-1994;   94US-0209762.
PR   11-JUL-1994;   94US-0273540.
PR   15-SEP-1994;   94US-0306473.
PR   13-JAN-1995;   95US-0372105.
PR   20-JUL-1995;   95US-0504841.
XX
PA   (XOMA ) XOMA CORP.
XX
PI   Lim E, Fadem MB, Little RG;
XX
DR   WPI; 2001-090160/10.
XX
PT   Novel anti-fungal peptides derived from domain III of
PT   bactericidal/permeability-increasing protein useful for killing or
PT   inhibiting replication of fungi and for treating fungal infections
XX
PS   Example 2; Columns 173-174; 134pp; English.
XX
CC   The present invention relates to antifungal peptides (see
CC   AAB65301-B65550) derived from Domain III (amino acids 142-169) of
CC   bactericidal/permeability-increasing protein (BPI). The present sequence
CC   is one such antifungal peptide. BPI is a protein isolated from the
CC   granules of mammalian polymorphonuclear leukocytes (PMNs or
CC   neutrophils). BPI has potent bactericidal activity against a broad range
CC   of gram-negative bacteria. The peptides of the present invention are
CC   useful for killing or inhibiting replication of fungi, and treating
CC   infections caused by fungus belonging to Candida, Aspergillus,
CC   Cryptococcus species such as C.albicans, C.glabrata, C.krusei,
CC   C.lusitaniae, C.parapsilosis and C.tropicalis.
XX
SQ   Sequence   9 AA;

  Query Match            100.0%;  Score 48;  DB 22;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 9.3e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


Search completed: October 14, 2003, 14:41:32

Job time : 38.7917 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 14:33:23 ; Search time 12.75 Seconds
(without alignments)
29.866 Million cell updates/sec

Title: US-09-881-490-164
Perfect score: 48
Sequence: 1 KKWLIQLKK 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	48	100.0	9	2	US-08-621-803-193	Sequence 193, App
2	48	100.0	9	2	US-08-621-803-197	Sequence 197, App
3	48	100.0	9	2	US-08-621-259A-160	Sequence 160, App
4	48	100.0	9	2	US-08-621-259A-162	Sequence 162, App
5	48	100.0	9	2	US-08-621-259A-164	Sequence 164, App
6	48	100.0	9	2	US-08-621-259A-173	Sequence 173, App
7	48	100.0	9	3	US-09-217-352-193	Sequence 193, App
8	48	100.0	9	3	US-09-217-352-197	Sequence 197, App
9	48	100.0	9	5	PCT-US95-09262-160	Sequence 160, App
10	48	100.0	9	5	PCT-US95-09262-162	Sequence 162, App
11	48	100.0	9	5	PCT-US95-09262-164	Sequence 164, App
12	48	100.0	9	5	PCT-US95-09262-173	Sequence 173, App
13	48	100.0	10	2	US-08-621-803-204	Sequence 204, App
14	48	100.0	10	2	US-08-621-803-205	Sequence 205, App
15	48	100.0	10	2	US-08-621-259A-179	Sequence 179, App
16	48	100.0	10	2	US-08-621-259A-180	Sequence 180, App
17	48	100.0	10	3	US-09-217-352-204	Sequence 204, App
18	48	100.0	10	3	US-09-217-352-205	Sequence 205, App
19	48	100.0	10	5	PCT-US95-09262-179	Sequence 179, App
20	48	100.0	10	5	PCT-US95-09262-180	Sequence 180, App
21	48	100.0	11	2	US-08-621-803-206	Sequence 206, App
22	48	100.0	11	2	US-08-621-259A-181	Sequence 181, App
23	48	100.0	11	3	US-09-217-352-206	Sequence 206, App
24	48	100.0	11	5	PCT-US95-09262-181	Sequence 181, App
25	43	89.6	8	2	US-08-621-803-192	Sequence 192, App
26	43	89.6	8	2	US-08-621-259A-157	Sequence 157, App
27	43	89.6	8	2	US-08-621-259A-159	Sequence 159, App
28	43	89.6	8	3	US-09-217-352-192	Sequence 192, App
29	43	89.6	8	5	PCT-US95-09262-157	Sequence 157, App
30	43	89.6	8	5	PCT-US95-09262-159	Sequence 159, App
31	43	89.6	9	2	US-08-621-803-198	Sequence 198, App
32	43	89.6	9	2	US-08-621-803-202	Sequence 202, App
33	43	89.6	9	2	US-08-621-803-203	Sequence 203, App
34	43	89.6	9	2	US-08-621-259A-165	Sequence 165, App
35	43	89.6	9	2	US-08-621-259A-166	Sequence 166, App
36	43	89.6	9	2	US-08-621-259A-175	Sequence 175, App
37	43	89.6	9	2	US-08-621-259A-176	Sequence 176, App
38	43	89.6	9	2	US-08-621-259A-206	Sequence 206, App
39	43	89.6	9	2	US-08-621-259A-241	Sequence 241, App
40	43	89.6	9	2	US-08-621-259A-242	Sequence 242, App
41	43	89.6	9	3	US-09-217-352-198	Sequence 198, App
42	43	89.6	9	3	US-09-217-352-202	Sequence 202, App
43	43	89.6	9	3	US-09-217-352-203	Sequence 203, App
44	43	89.6	9	5	PCT-US95-09262-165	Sequence 165, App
45	43	89.6	9	5	PCT-US95-09262-166	Sequence 166, App

ALIGNMENTS

```
RESULT 1
US-08-621-803-193
; Sequence 193, Application US/08621803
; Patent No. 5851802
;  GENERAL INFORMATION:
;    APPLICANT:  Better, Marc D.
;    TITLE OF INVENTION:  Methods for Recombinant Microbial Production of
;    TITLE OF INVENTION:  Fusion Proteins and BPI-Derived Peptides
;    NUMBER OF SEQUENCES:  265
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,803
;      FILING DATE:  22-MAR-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25,447
;      REFERENCE/DOCKET NUMBER:  27129/33199
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  193:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.327"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-08-621-803-193

  Query Match             100.0%;  Score 48;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.5e+05;
```

```
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


RESULT 2
US-08-621-803-197
; Sequence 197, Application US/08621803
; Patent No. 5851802
;  GENERAL INFORMATION:
;    APPLICANT:  Better, Marc D.
;    TITLE OF INVENTION:  Methods for Recombinant Microbial Production of
;    TITLE OF INVENTION:  Fusion Proteins and BPI-Derived Peptides
;    NUMBER OF SEQUENCES:  265
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,803
;      FILING DATE:  22-MAR-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25,447
;      REFERENCE/DOCKET NUMBER:  27129/33199
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  197:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.331"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1
;      OTHER INFORMATION:  /label= Acetylated
;      OTHER INFORMATION:  /note= "Position 1 is acetylated."
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-08-621-803-197

  Query Match             100.0%;  Score 48;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.5e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


RESULT 3
US-08-621-259A-160
; Sequence 160, Application US/08621259A
; Patent No. 5858974
;  GENERAL INFORMATION:
;    APPLICANT:  Little II, Roger G
;    APPLICANT:  Lim, Edward
;    APPLICANT:  Fadem, Mitchell B.
;    TITLE OF INVENTION:  Anti-Fungal Peptides
;    NUMBER OF SEQUENCES:  252
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McAndrews, Held & Malloy, Ltd.
;      STREET:  500 West Madison Street
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60661
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,259A
;      FILING DATE:  21-MAR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/504,841
;      FILING DATE:  20-JUL-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McNicholas, Janet M.
;      REGISTRATION NUMBER:  32,918
;      REFERENCE/DOCKET NUMBER:  11021US02
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/707-8889
;      TELEFAX:  312/707-9155
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  160:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.327"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-08-621-259A-160

  Query Match             100.0%;  Score 48;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.5e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


RESULT 4
US-08-621-259A-162
; Sequence 162, Application US/08621259A
; Patent No. 5858974
;  GENERAL INFORMATION:
;    APPLICANT:  Little II, Roger G
;    APPLICANT:  Lim, Edward
;    APPLICANT:  Fadem, Mitchell B.
;    TITLE OF INVENTION:  Anti-Fungal Peptides
;    NUMBER OF SEQUENCES:  252
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McAndrews, Held & Malloy, Ltd.
;      STREET:  500 West Madison Street
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
```

```
;       ZIP:  60661
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/621,259A
;       FILING DATE:  21-MAR-1996
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/504,841
;       FILING DATE:  20-JUL-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  McNicholas, Janet M.
;       REGISTRATION NUMBER:  32,918
;       REFERENCE/DOCKET NUMBER:  11021US02
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  312/707-8889
;       TELEFAX:  312/707-9155
;       TELEX:
;   INFORMATION FOR SEQ ID NO:  162:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  9 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
;     FEATURE:
;       NAME/KEY:  misc_feature
;       OTHER INFORMATION:  "XMP.331"
;     FEATURE:
;       NAME/KEY:  Modified-site
;       LOCATION:  1
;       OTHER INFORMATION:  /label= Acetylated
;       OTHER INFORMATION:  /note= "Position 1 is acetylated."
;     FEATURE:
;       NAME/KEY:  Modified-site
;       LOCATION:  C-Terminus
;       OTHER INFORMATION:  /label= Amidation
;       OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-08-621-259A-162

  Query Match             100.0%;  Score 48;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


RESULT 5
US-08-621-259A-164
; Sequence 164, Application US/08621259A
; Patent No. 5858974
;  GENERAL INFORMATION:
;    APPLICANT:  Little II, Roger G
;    APPLICANT:  Lim, Edward
;    APPLICANT:  Fadem, Mitchell B.
;    TITLE OF INVENTION:  Anti-Fungal Peptides
;    NUMBER OF SEQUENCES:  252
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McAndrews, Held & Malloy, Ltd.
;      STREET:  500 West Madison Street
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60661
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,259A
;      FILING DATE:  21-MAR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/504,841
;      FILING DATE:  20-JUL-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McNicholas, Janet M.
;      REGISTRATION NUMBER:  32,918
;      REFERENCE/DOCKET NUMBER:  11021US02
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/707-8889
;      TELEFAX:  312/707-9155
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  164:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.333"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1
;      OTHER INFORMATION:  /label= D-Lys
;      OTHER INFORMATION:  /note= "Position 1 is D-lysine"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-08-621-259A-164

  Query Match             100.0%;  Score 48;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.5e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


RESULT 6
US-08-621-259A-173
; Sequence 173, Application US/08621259A
; Patent No. 5858974
;  GENERAL INFORMATION:
;    APPLICANT:  Little II, Roger G
;    APPLICANT:  Lim, Edward
;    APPLICANT:  Fadem, Mitchell B.
;    TITLE OF INVENTION:  Anti-Fungal Peptides
;    NUMBER OF SEQUENCES:  252
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McAndrews, Held & Malloy, Ltd.
;      STREET:  500 West Madison Street
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60661
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,259A
;      FILING DATE:  21-MAR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/504,841
;      FILING DATE:  20-JUL-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McNicholas, Janet M.
```

```
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.331"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1
;      OTHER INFORMATION:  /label= Acetylated
;      OTHER INFORMATION:  /note= "Position 1 is acetylated"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated"
PCT-US95-09262-162

  Query Match             100.0%;  Score 48;  DB 5;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


RESULT 11
PCT-US95-09262-164
; Sequence 164, Application PC/TUS9509262
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  Anti-Fungal Peptides
;    NUMBER OF SEQUENCES:  206
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/09262
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/372,105
;      FILING DATE:  13-JAN-95
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/306,473
;      FILING DATE:  15-SEP-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/273,540
;      FILING DATE:  11-JUL-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/209,762
;      FILING DATE:  11-MAR-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/183,222
;      FILING DATE:  14-JAN-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/093,202
;      FILING DATE:  15-JUL-93
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/030,644
;      FILING DATE:  12-MAR-93
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25,447
;      REFERENCE/DOCKET NUMBER:  27129/10040
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  164:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.333"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1
;      OTHER INFORMATION:  /label= D-Lys
;      OTHER INFORMATION:  /note= "Position 1 is D-lysine"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated"
PCT-US95-09262-164

  Query Match             100.0%;  Score 48;  DB 5;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.5e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


RESULT 12
PCT-US95-09262-173
; Sequence 173, Application PC/TUS9509262
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  Anti-Fungal Peptides
;    NUMBER OF SEQUENCES:  206
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/09262
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/372,105
;      FILING DATE:  13-JAN-95
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/306,473
;      FILING DATE:  15-SEP-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/273,540
;      FILING DATE:  11-JUL-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/209,762
;      FILING DATE:  11-MAR-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/183,222
;      FILING DATE:  14-JAN-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/093,202
```

```
;      FILING DATE:  15-JUL-93
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/030,644
;      FILING DATE:  12-MAR-93
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25,447
;      REFERENCE/DOCKET NUMBER:  27129/10040
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  173:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.342"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1, 2, 8 & 9
;      OTHER INFORMATION:  /label= D-Lys
;      OTHER INFORMATION:  /note= "Positions 1, 2, 8 & 9 are D-lysine"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated"
PCT-US95-09262-173

  Query Match             100.0%;  Score 48;  DB 5;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.5e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


RESULT 13
US-08-621-803-204
; Sequence 204, Application US/08621803
; Patent No. 5851802
;  GENERAL INFORMATION:
;    APPLICANT:  Better, Marc D.
;    TITLE OF INVENTION:  Methods for Recombinant Microbial Production of
;    TITLE OF INVENTION:  Fusion Proteins and BPI-Derived Peptides
;    NUMBER OF SEQUENCES:  265
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,803
;      FILING DATE:  22-MAR-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25,447
;      REFERENCE/DOCKET NUMBER:  27129/33199
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  204:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.348"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-08-621-803-204

  Query Match             100.0%;  Score 48;  DB 2;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 0.24;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          2 KKWLIQLKK 10


RESULT 14
US-08-621-803-205
; Sequence 205, Application US/08621803
; Patent No. 5851802
;  GENERAL INFORMATION:
;    APPLICANT:  Better, Marc D.
;    TITLE OF INVENTION:  Methods for Recombinant Microbial Production of
;    TITLE OF INVENTION:  Fusion Proteins and BPI-Derived Peptides
;    NUMBER OF SEQUENCES:  265
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,803
;      FILING DATE:  22-MAR-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25,447
;      REFERENCE/DOCKET NUMBER:  27129/33199
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  205:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.349"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-08-621-803-205
```

```
  Query Match           100.0%;  Score 48;  DB 2;  Length 10;
  Best Local Similarity  100.0%;  Pred. No. 0.24;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


RESULT 15
US-08-621-259A-179
; Sequence 179, Application US/08621259A
; Patent No. 5858974
;  GENERAL INFORMATION:
;    APPLICANT:  Little II, Roger G
;    APPLICANT:  Lim, Edward
;    APPLICANT:  Fadem, Mitchell B.
;    TITLE OF INVENTION:  Anti-Fungal Peptides
;    NUMBER OF SEQUENCES:  252
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McAndrews, Held & Malloy, Ltd.
;      STREET:  500 West Madison Street
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60661
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,259A
;      FILING DATE:  21-MAR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/504,841
;      FILING DATE:  20-JUL-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McNicholas, Janet M.
;      REGISTRATION NUMBER:  32,918
;      REFERENCE/DOCKET NUMBER:  11021US02
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/707-8889
;      TELEFAX:  312/707-9155
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  179:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.348"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-08-621-259A-179

  Query Match           100.0%;  Score 48;  DB 2;  Length 10;
  Best Local Similarity  100.0%;  Pred. No. 0.24;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          2 KKWLIQLKK 10


Search completed: October 14, 2003, 14:42:46
Job time : 12.75 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 14:34:14 ; Search time 21.9375 Seconds
(without alignments)
66.104 Million cell updates/sec

Title: US-09-881-490-164
Perfect score: 48
Sequence: 1 KKWLIQLKK 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 600653 seqs, 161128416 residues

Total number of hits satisfying chosen parameters: 600653

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	48	100.0	9	9	US-09-765-527-193	Sequence 193, App
2	48	100.0	9	9	US-09-765-527-197	Sequence 197, App
3	48	100.0	9	9	US-09-881-490-160	Sequence 160, App
4	48	100.0	9	9	US-09-881-490-162	Sequence 162, App
5	48	100.0	9	9	US-09-881-490-164	Sequence 164, App
6	48	100.0	9	9	US-09-881-490-173	Sequence 173, App
7	48	100.0	10	9	US-09-765-527-204	Sequence 204, App
8	48	100.0	10	9	US-09-765-527-205	Sequence 205, App
9	48	100.0	10	9	US-09-881-490-179	Sequence 179, App
10	48	100.0	10	9	US-09-881-490-180	Sequence 180, App
11	48	100.0	11	9	US-09-765-527-206	Sequence 206, App
12	48	100.0	11	9	US-09-881-490-181	Sequence 181, App
13	43	89.6	8	9	US-09-765-527-192	Sequence 192, App
14	43	89.6	8	9	US-09-881-490-157	Sequence 157, App
15	43	89.6	8	9	US-09-881-490-159	Sequence 159, App
16	43	89.6	9	9	US-09-765-527-198	Sequence 198, App
17	43	89.6	9	9	US-09-765-527-202	Sequence 202, App
18	43	89.6	9	9	US-09-765-527-203	Sequence 203, App
19	43	89.6	9	9	US-09-881-490-165	Sequence 165, App
20	43	89.6	9	9	US-09-881-490-166	Sequence 166, App
21	43	89.6	9	9	US-09-881-490-175	Sequence 175, App
22	43	89.6	9	9	US-09-881-490-176	Sequence 176, App
23	43	89.6	10	9	US-09-765-527-216	Sequence 216, App
24	40	83.3	10	9	US-09-765-527-207	Sequence 207, App
25	40	83.3	10	9	US-09-881-490-182	Sequence 182, App
26	38	79.2	7	9	US-09-765-527-188	Sequence 188, App
27	38	79.2	7	9	US-09-765-527-189	Sequence 189, App
28	38	79.2	7	9	US-09-765-527-190	Sequence 190, App
29	38	79.2	7	9	US-09-765-527-191	Sequence 191, App
30	38	79.2	7	9	US-09-881-490-155	Sequence 155, App
31	38	79.2	7	9	US-09-881-490-156	Sequence 156, App
32	38	79.2	7	9	US-09-881-490-158	Sequence 158, App
33	38	79.2	9	9	US-09-765-527-236	Sequence 236, App
34	38	79.2	9	9	US-09-881-490-177	Sequence 177, App
35	38	79.2	10	9	US-09-765-527-214	Sequence 214, App
36	38	79.2	11	9	US-09-765-527-208	Sequence 208, App
37	38	79.2	11	9	US-09-881-490-183	Sequence 183, App
38	38	79.2	11	9	US-09-881-490-189	Sequence 189, App
39	37	77.1	10	9	US-09-765-527-218	Sequence 218, App
40	36	75.0	9	9	US-09-765-527-199	Sequence 199, App
41	36	75.0	9	9	US-09-881-490-167	Sequence 167, App
42	36	75.0	436	10	US-09-925-300-1640	Sequence 1640, Ap
43	35	72.9	55	11	US-09-764-891-3233	Sequence 3233, Ap
44	35	72.9	2073	11	US-09-978-244A-10	Sequence 10, Appl
45	34	70.8	9	9	US-09-765-527-201	Sequence 201, App

ALIGNMENTS

```
RESULT 1
US-09-765-527-193
; Sequence 193, Application US/09765527
; Patent No. US20020006638A1
;   GENERAL INFORMATION:
;        APPLICANT: Better, Marc D.
;        TITLE OF INVENTION: Methods for Recombinant Microbial Production of
;                            Fusion Proteins and BPI-Derived Peptides
;        NUMBER OF SEQUENCES: 265
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
;             STREET: 6300 Sears Tower, 233 South Wacker Drive
;             CITY: Chicago
;             STATE: Illinois
;             COUNTRY: United States of America
;             ZIP: 60606-6402
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/765,527
;             FILING DATE: 18-Jan-2001
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/621,803
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Borun, Michael F.
;             REGISTRATION NUMBER: 25,447
;             REFERENCE/DOCKET NUMBER: 27129/33199
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 312/474-6300
;             TELEFAX: 312/474-0448
;             TELEX: 25-3856
;   INFORMATION FOR SEQ ID NO: 193:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 9 amino acids
```

```
;             TYPE: amino acid
;             TOPOLOGY: linear
;          MOLECULE TYPE: peptide
;          FEATURE:
;             NAME/KEY:  misc_feature
;             OTHER INFORMATION:  "XMP.327"
;          FEATURE:
;             NAME/KEY:  Modified-site
;             LOCATION:  C-Terminus
;             OTHER INFORMATION:  /label= Amidation
;             /note= "The C-Terminus is Amidated."
;          SEQUENCE DESCRIPTION: SEQ ID NO: 193:
US-09-765-527-193

  Query Match             100.0%;  Score 48;  DB 9;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 5.4e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


RESULT 2
US-09-765-527-197
; Sequence 197, Application US/09765527
; Patent No. US20020006638A1
;   GENERAL INFORMATION:
;        APPLICANT: Better, Marc D.
;        TITLE OF INVENTION: Methods for Recombinant Microbial Production of
;                            Fusion Proteins and BPI-Derived Peptides
;        NUMBER OF SEQUENCES: 265
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
;             STREET: 6300 Sears Tower, 233 South Wacker Drive
;             CITY: Chicago
;             STATE: Illinois
;             COUNTRY: United States of America
;             ZIP: 60606-6402
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/765,527
;             FILING DATE: 18-Jan-2001
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/621,803
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Borun, Michael F.
;             REGISTRATION NUMBER: 25,447
;             REFERENCE/DOCKET NUMBER: 27129/33199
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 312/474-6300
;             TELEFAX: 312/474-0448
;             TELEX: 25-3856
;   INFORMATION FOR SEQ ID NO: 197:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 9 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        FEATURE:
;             NAME/KEY:  misc_feature
;             OTHER INFORMATION:  "XMP.331"
;        FEATURE:
;             NAME/KEY:  Modified-site
;             LOCATION:  1
;             OTHER INFORMATION:  /label= Acetylated
;             /note= "Position 1 is acetylated."
;        FEATURE:
;             NAME/KEY:  Modified-site
;             LOCATION:  C-Terminus
;             OTHER INFORMATION:  /label= Amidation
;             /note= "The C-Terminus is Amidated."
;          SEQUENCE DESCRIPTION: SEQ ID NO: 197:
US-09-765-527-197

  Query Match             100.0%;  Score 48;  DB 9;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 5.4e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


RESULT 3
US-09-881-490-160
; Sequence 160, Application US/09881490
; Patent No. US20020077298A1
;   GENERAL INFORMATION:
;        APPLICANT: Little II, Roger G.
;                   Lim, Edward
;                   Fadem, Mitchell B.
;        TITLE OF INVENTION: Anti-Fungal Peptides
;        NUMBER OF SEQUENCES: 211
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: McAndrews, Held & Malloy, Ltd.
;             STREET: 500 West Madison Street, 34th FloorDrive
;             CITY: Chicago
;             STATE: Illinois
;             COUNTRY: United States of America
;             ZIP: 60661
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/881,490
;             FILING DATE: 14-Jun-2001
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/119,858
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: 08/372,105
;             FILING DATE: 13-JAN-95
;             APPLICATION NUMBER: 08/306,473
;             FILING DATE: 15-SEP-94
;             APPLICATION NUMBER: 08/273,540
;             FILING DATE: 11-JUL-94
;             APPLICATION NUMBER: 08/209,762
;             FILING DATE: 11-MAR-94
;             APPLICATION NUMBER: 08/183,222
;             FILING DATE: 14-JAN-94
;             APPLICATION NUMBER: 08/093,202
;             FILING DATE: 15-JUL-93
;             APPLICATION NUMBER: 08/030,644
;             FILING DATE: 12-MAR-93
;        ATTORNEY/AGENT INFORMATION:
;             NAME: McNicholas, Janet M.
;             REGISTRATION NUMBER: 32,918
;             REFERENCE/DOCKET NUMBER: 100-238/11021US01
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 312/707-8889
;             TELEFAX: 312/707-9155
;             TELEX: 650 388-1248
;   INFORMATION FOR SEQ ID NO: 160:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 9 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        FEATURE:
```

```
;             NAME/KEY:  misc_feature
;             OTHER INFORMATION:  "XMP.327"
;        FEATURE:
;             NAME/KEY:  Modified-site
;             LOCATION:  C-Terminus
;             OTHER INFORMATION:  /label= Amidation
;             /note= "The C-Terminus is Amidated"
;        SEQUENCE DESCRIPTION: SEQ ID NO: 160:
US-09-881-490-160

  Query Match             100.0%;  Score 48;  DB 9;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 5.4e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


RESULT 4
US-09-881-490-162
; Sequence 162, Application US/09881490
; Patent No. US20020077298A1
;   GENERAL INFORMATION:
;        APPLICANT: Little II, Roger G.
;                   Lim, Edward
;                   Fadem, Mitchell B.
;        TITLE OF INVENTION: Anti-Fungal Peptides
;        NUMBER OF SEQUENCES: 211
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: McAndrews, Held & Malloy, Ltd.
;             STREET: 500 West Madison Street, 34th FloorDrive
;             CITY: Chicago
;             STATE: Illinois
;             COUNTRY: United States of America
;             ZIP: 60661
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/881,490
;             FILING DATE: 14-Jun-2001
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/119,858
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: 08/372,105
;             FILING DATE: 13-JAN-95
;             APPLICATION NUMBER: 08/306,473
;             FILING DATE: 15-SEP-94
;             APPLICATION NUMBER: 08/273,540
;             FILING DATE: 11-JUL-94
;             APPLICATION NUMBER: 08/209,762
;             FILING DATE: 11-MAR-94
;             APPLICATION NUMBER: 08/183,222
;             FILING DATE: 14-JAN-94
;             APPLICATION NUMBER: 08/093,202
;             FILING DATE: 15-JUL-93
;             APPLICATION NUMBER: 08/030,644
;             FILING DATE: 12-MAR-93
;        ATTORNEY/AGENT INFORMATION:
;             NAME: McNicholas, Janet M.
;             REGISTRATION NUMBER: 32,918
;             REFERENCE/DOCKET NUMBER: 100-238/11021US01
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 312/707-8889
;             TELEFAX: 312/707-9155
;             TELEX: 650 388-1248
;   INFORMATION FOR SEQ ID NO: 162:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 9 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        FEATURE:
;             NAME/KEY:  misc_feature
;             OTHER INFORMATION:  "XMP.331"
;        FEATURE:
;             NAME/KEY:  Modified-site
;             LOCATION:  1
;             OTHER INFORMATION:  /label= Acetylated
;             /note= "Position 1 is acetylated"
;        FEATURE:
;             NAME/KEY:  Modified-site
;             LOCATION:  C-Terminus
;             OTHER INFORMATION:  /label= Amidation
;             /note= "The C-Terminus is Amidated"
;        SEQUENCE DESCRIPTION: SEQ ID NO: 162:
US-09-881-490-162

  Query Match             100.0%;  Score 48;  DB 9;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 5.4e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


RESULT 5
US-09-881-490-164
; Sequence 164, Application US/09881490
; Patent No. US20020077298A1
;   GENERAL INFORMATION:
;        APPLICANT: Little II, Roger G.
;                   Lim, Edward
;                   Fadem, Mitchell B.
;        TITLE OF INVENTION: Anti-Fungal Peptides
;        NUMBER OF SEQUENCES: 211
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: McAndrews, Held & Malloy, Ltd.
;             STREET: 500 West Madison Street, 34th FloorDrive
;             CITY: Chicago
;             STATE: Illinois
;             COUNTRY: United States of America
;             ZIP: 60661
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/881,490
;             FILING DATE: 14-Jun-2001
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/119,858
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: 08/372,105
;             FILING DATE: 13-JAN-95
;             APPLICATION NUMBER: 08/306,473
;             FILING DATE: 15-SEP-94
;             APPLICATION NUMBER: 08/273,540
;             FILING DATE: 11-JUL-94
;             APPLICATION NUMBER: 08/209,762
;             FILING DATE: 11-MAR-94
;             APPLICATION NUMBER: 08/183,222
;             FILING DATE: 14-JAN-94
;             APPLICATION NUMBER: 08/093,202
;             FILING DATE: 15-JUL-93
;             APPLICATION NUMBER: 08/030,644
;             FILING DATE: 12-MAR-93
;        ATTORNEY/AGENT INFORMATION:
;             NAME: McNicholas, Janet M.
;             REGISTRATION NUMBER: 32,918
;             REFERENCE/DOCKET NUMBER: 100-238/11021US01
```

```
;         TELECOMMUNICATION INFORMATION:
;              TELEPHONE: 312/707-8889
;              TELEFAX: 312/707-9155
;              TELEX: 650 388-1248
;    INFORMATION FOR SEQ ID NO: 164:
;         SEQUENCE CHARACTERISTICS:
;              LENGTH: 9 amino acids
;              TYPE: amino acid
;              TOPOLOGY: linear
;         MOLECULE TYPE: peptide
;         FEATURE:
;              NAME/KEY:  misc_feature
;              OTHER INFORMATION:  "XMP.333"
;         FEATURE:
;              NAME/KEY:  Modified-site
;              LOCATION:  1
;              OTHER INFORMATION:  /label= D-Lys
;              /note= "Position 1 is D-lysine"
;         FEATURE:
;              NAME/KEY:  Modified-site
;              LOCATION:  C-Terminus
;              OTHER INFORMATION:  /label= Amidation
;              /note= "The C-Terminus is Amidated"
;         SEQUENCE DESCRIPTION: SEQ ID NO: 164:
US-09-881-490-164

  Query Match             100.0%;  Score 48;  DB 9;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 5.4e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


RESULT 6
US-09-881-490-173
; Sequence 173, Application US/09881490
; Patent No. US20020077298A1
;   GENERAL INFORMATION:
;         APPLICANT: Little II, Roger G.
;                    Lim, Edward
;                    Fadem, Mitchell B.
;         TITLE OF INVENTION: Anti-Fungal Peptides
;         NUMBER OF SEQUENCES: 211
;         CORRESPONDENCE ADDRESS:
;              ADDRESSEE: McAndrews, Held & Malloy, Ltd.
;              STREET: 500 West Madison Street, 34th FloorDrive
;              CITY: Chicago
;              STATE: Illinois
;              COUNTRY: United States of America
;              ZIP: 60661
;         COMPUTER READABLE FORM:
;              MEDIUM TYPE: Floppy disk
;              COMPUTER: IBM PC compatible
;              OPERATING SYSTEM: PC-DOS/MS-DOS
;              SOFTWARE: PatentIn Release #1.0, Version #1.25
;         CURRENT APPLICATION DATA:
;              APPLICATION NUMBER: US/09/881,490
;              FILING DATE: 14-Jun-2001
;         PRIOR APPLICATION DATA:
;              APPLICATION NUMBER: 09/119,858
;              FILING DATE: <Unknown>
;              APPLICATION NUMBER: 08/372,105
;              FILING DATE: 13-JAN-95
;              APPLICATION NUMBER: 08/306,473
;              FILING DATE: 15-SEP-94
;              APPLICATION NUMBER: 08/273,540
;              FILING DATE: 11-JUL-94
;              APPLICATION NUMBER: 08/209,762
;              FILING DATE: 11-MAR-94
;              APPLICATION NUMBER: 08/183,222
;              FILING DATE: 14-JAN-94
;              APPLICATION NUMBER: 08/093,202
;              FILING DATE: 15-JUL-93
;              APPLICATION NUMBER: 08/030,644
;              FILING DATE: 12-MAR-93
;         ATTORNEY/AGENT INFORMATION:
;              NAME: McNicholas, Janet M.
;              REGISTRATION NUMBER: 32,918
;              REFERENCE/DOCKET NUMBER: 100-238/11021US01
;         TELECOMMUNICATION INFORMATION:
;              TELEPHONE: 312/707-8889
;              TELEFAX: 312/707-9155
;              TELEX: 650 388-1248
;    INFORMATION FOR SEQ ID NO: 173:
;         SEQUENCE CHARACTERISTICS:
;              LENGTH: 9 amino acids
;              TYPE: amino acid
;              TOPOLOGY: linear
;         MOLECULE TYPE: peptide
;         FEATURE:
;              NAME/KEY:  misc_feature
;              OTHER INFORMATION:  "XMP.342"
;         FEATURE:
;              NAME/KEY:  Modified-site
;              LOCATION:  1, 2, 8 & 9
;              OTHER INFORMATION:  /label= D-Lys
;              /note= "Positions 1, 2, 8 & 9 are D-lysine"
;         FEATURE:
;              NAME/KEY:  Modified-site
;              LOCATION:  C-Terminus
;              OTHER INFORMATION:  /label= Amidation
;              /note= "The C-Terminus is Amidated"
;         SEQUENCE DESCRIPTION: SEQ ID NO: 173:
US-09-881-490-173

  Query Match             100.0%;  Score 48;  DB 9;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 5.4e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


RESULT 7
US-09-765-527-204
; Sequence 204, Application US/09765527
; Patent No. US20020006638A1
;   GENERAL INFORMATION:
;         APPLICANT: Better, Marc D.
;         TITLE OF INVENTION: Methods for Recombinant Microbial Production of
;                             Fusion Proteins and BPI-Derived Peptides
;         NUMBER OF SEQUENCES: 265
;         CORRESPONDENCE ADDRESS:
;              ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
;              STREET: 6300 Sears Tower, 233 South Wacker Drive
;              CITY: Chicago
;              STATE: Illinois
;              COUNTRY: United States of America
;              ZIP: 60606-6402
;         COMPUTER READABLE FORM:
;              MEDIUM TYPE: Floppy disk
;              COMPUTER: IBM PC compatible
;              OPERATING SYSTEM: PC-DOS/MS-DOS
;              SOFTWARE: PatentIn Release #1.0, Version #1.25
;         CURRENT APPLICATION DATA:
;              APPLICATION NUMBER: US/09/765,527
;              FILING DATE: 18-Jan-2001
;         PRIOR APPLICATION DATA:
;              APPLICATION NUMBER: 08/621,803
;              FILING DATE: <Unknown>
;         ATTORNEY/AGENT INFORMATION:
;              NAME: Borun, Michael F.
;              REGISTRATION NUMBER: 25,447
```

```
;             REFERENCE/DOCKET NUMBER: 27129/33199
;          TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 312/474-6300
;             TELEFAX: 312/474-0448
;             TELEX: 25-3856
;       INFORMATION FOR SEQ ID NO: 204:
;          SEQUENCE CHARACTERISTICS:
;             LENGTH: 10 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;          MOLECULE TYPE: peptide
;          FEATURE:
;             NAME/KEY:  misc_feature
;             OTHER INFORMATION:  "XMP.348"
;          FEATURE:
;             NAME/KEY:  Modified-site
;             LOCATION:  C-Terminus
;             OTHER INFORMATION:  /label= Amidation
;             /note= "The C-Terminus is Amidated."
;          SEQUENCE DESCRIPTION: SEQ ID NO: 204:
US-09-765-527-204

  Query Match             100.0%;  Score 48;  DB 9;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 0.13;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          2 KKWLIQLKK 10


RESULT 8
US-09-765-527-205
; Sequence 205, Application US/09765527
; Patent No. US20020006638A1
;   GENERAL INFORMATION:
;          APPLICANT: Better, Marc D.
;          TITLE OF INVENTION: Methods for Recombinant Microbial Production of
;                              Fusion Proteins and BPI-Derived Peptides
;          NUMBER OF SEQUENCES: 265
;          CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
;             STREET: 6300 Sears Tower, 233 South Wacker Drive
;             CITY: Chicago
;             STATE: Illinois
;             COUNTRY: United States of America
;             ZIP: 60606-6402
;          COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;          CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/765,527
;             FILING DATE: 18-Jan-2001
;          PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/621,803
;             FILING DATE: <Unknown>
;          ATTORNEY/AGENT INFORMATION:
;             NAME: Borun, Michael F.
;             REGISTRATION NUMBER: 25,447
;             REFERENCE/DOCKET NUMBER: 27129/33199
;          TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 312/474-6300
;             TELEFAX: 312/474-0448
;             TELEX: 25-3856
;       INFORMATION FOR SEQ ID NO: 205:
;          SEQUENCE CHARACTERISTICS:
;             LENGTH: 10 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;          MOLECULE TYPE: peptide
;          FEATURE:
;             NAME/KEY:  misc_feature
;             OTHER INFORMATION:  "XMP.349"
;          FEATURE:
;             NAME/KEY:  Modified-site
;             LOCATION:  C-Terminus
;             OTHER INFORMATION:  /label= Amidation
;             /note= "The C-Terminus is Amidated."
;          SEQUENCE DESCRIPTION: SEQ ID NO: 205:
US-09-765-527-205

  Query Match             100.0%;  Score 48;  DB 9;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 0.13;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


RESULT 9
US-09-881-490-179
; Sequence 179, Application US/09881490
; Patent No. US20020077298A1
;   GENERAL INFORMATION:
;          APPLICANT: Little II, Roger G.
;                     Lim, Edward
;                     Fadem, Mitchell B.
;          TITLE OF INVENTION: Anti-Fungal Peptides
;          NUMBER OF SEQUENCES: 211
;          CORRESPONDENCE ADDRESS:
;             ADDRESSEE: McAndrews, Held & Malloy, Ltd.
;             STREET: 500 West Madison Street, 34th FloorDrive
;             CITY: Chicago
;             STATE: Illinois
;             COUNTRY: United States of America
;             ZIP: 60661
;          COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;          CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/881,490
;             FILING DATE: 14-Jun-2001
;          PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/119,858
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: 08/372,105
;             FILING DATE: 13-JAN-95
;             APPLICATION NUMBER: 08/306,473
;             FILING DATE: 15-SEP-94
;             APPLICATION NUMBER: 08/273,540
;             FILING DATE: 11-JUL-94
;             APPLICATION NUMBER: 08/209,762
;             FILING DATE: 11-MAR-94
;             APPLICATION NUMBER: 08/183,222
;             FILING DATE: 14-JAN-94
;             APPLICATION NUMBER: 08/093,202
;             FILING DATE: 15-JUL-93
;             APPLICATION NUMBER: 08/030,644
;             FILING DATE: 12-MAR-93
;          ATTORNEY/AGENT INFORMATION:
;             NAME: McNicholas, Janet M.
;             REGISTRATION NUMBER: 32,918
;             REFERENCE/DOCKET NUMBER: 100-238/11C21US01
;          TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 312/707-8889
;             TELEFAX: 312/707-9155
;             TELEX: 650 388-1248
;       INFORMATION FOR SEQ ID NO: 179:
;          SEQUENCE CHARACTERISTICS:
;             LENGTH: 10 amino acids
;             TYPE: amino acid
```

```
;           TOPOLOGY: linear
;       MOLECULE TYPE: peptide
;       FEATURE:
;           NAME/KEY:  misc_feature
;           OTHER INFORMATION:  "XMP.348"
;       FEATURE:
;           NAME/KEY:  Modified-site
;           LOCATION:  C-Terminus
;           OTHER INFORMATION:  /label= Amidation
;           /note= "The C-Terminus is Amidated"
;       SEQUENCE DESCRIPTION: SEQ ID NO: 179:
US-09-881-490-179

  Query Match             100.0%;  Score 48;  DB 9;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 0.13;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          2 KKWLIQLKK 10


RESULT 10
US-09-881-490-180
; Sequence 180, Application US/09881490
; Patent No. US20020077298A1
;   GENERAL INFORMATION:
;       APPLICANT: Little II, Roger G.
;                  Lim, Edward
;                  Fadem, Mitchell B.
;       TITLE OF INVENTION: Anti-Fungal Peptides
;       NUMBER OF SEQUENCES: 211
;       CORRESPONDENCE ADDRESS:
;           ADDRESSEE: McAndrews, Held & Malloy, Ltd.
;           STREET: 500 West Madison Street, 34th FloorDrive
;           CITY: Chicago
;           STATE: Illinois
;           COUNTRY: United States of America
;           ZIP: 60661
;       COMPUTER READABLE FORM:
;           MEDIUM TYPE: Floppy disk
;           COMPUTER: IBM PC compatible
;           OPERATING SYSTEM: PC-DOS/MS-DOS
;           SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;           APPLICATION NUMBER: US/09/881,490
;           FILING DATE: 14-Jun-2001
;       PRIOR APPLICATION DATA:
;           APPLICATION NUMBER: 09/119,858
;           FILING DATE: <Unknown>
;           APPLICATION NUMBER: 08/372,105
;           FILING DATE: 13-JAN-95
;           APPLICATION NUMBER: 08/306,473
;           FILING DATE: 15-SEP-94
;           APPLICATION NUMBER: 08/273,540
;           FILING DATE: 11-JUL-94
;           APPLICATION NUMBER: 08/209,762
;           FILING DATE: 11-MAR-94
;           APPLICATION NUMBER: 08/183,222
;           FILING DATE: 14-JAN-94
;           APPLICATION NUMBER: 08/093,202
;           FILING DATE: 15-JUL-93
;           APPLICATION NUMBER: 08/030,644
;           FILING DATE: 12-MAR-93
;       ATTORNEY/AGENT INFORMATION:
;           NAME: McNicholas, Janet M.
;           REGISTRATION NUMBER: 32,918
;           REFERENCE/DOCKET NUMBER: 100-238/11021US01
;       TELECOMMUNICATION INFORMATION:
;           TELEPHONE: 312/707-8889
;           TELEFAX: 312/707-9155
;           TELEX: 650 388-1248
;   INFORMATION FOR SEQ ID NO: 180:
;       SEQUENCE CHARACTERISTICS:
;           LENGTH: 10 amino acids
;           TYPE: amino acid
;           TOPOLOGY: linear
;       MOLECULE TYPE: peptide
;       FEATURE:
;           NAME/KEY:  misc_feature
;           OTHER INFORMATION:  "XMP.349"
;       FEATURE:
;           NAME/KEY:  Modified-site
;           LOCATION:  C-Terminus
;           OTHER INFORMATION:  /label= Amidation
;           /note= "The C-Terminus is Amidated"
;       SEQUENCE DESCRIPTION: SEQ ID NO: 180:
US-09-881-490-180

  Query Match             100.0%;  Score 48;  DB 9;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 0.13;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          1 KKWLIQLKK 9


RESULT 11
US-09-765-527-206
; Sequence 206, Application US/09765527
; Patent No. US20020006638A1
;   GENERAL INFORMATION:
;       APPLICANT: Better, Marc D.
;       TITLE OF INVENTION: Methods for Recombinant Microbial Production of
;                           Fusion Proteins and BPI-Derived Peptides
;       NUMBER OF SEQUENCES: 265
;       CORRESPONDENCE ADDRESS:
;           ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
;           STREET: 6300 Sears Tower, 233 South Wacker Drive
;           CITY: Chicago
;           STATE: Illinois
;           COUNTRY: United States of America
;           ZIP: 60606-6402
;       COMPUTER READABLE FORM:
;           MEDIUM TYPE: Floppy disk
;           COMPUTER: IBM PC compatible
;           OPERATING SYSTEM: PC-DOS/MS-DOS
;           SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;           APPLICATION NUMBER: US/09/765,527
;           FILING DATE: 18-Jan-2001
;       PRIOR APPLICATION DATA:
;           APPLICATION NUMBER: 08/621,803
;           FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;           NAME: Borun, Michael F.
;           REGISTRATION NUMBER: 25,447
;           REFERENCE/DOCKET NUMBER: 27129/33199
;       TELECOMMUNICATION INFORMATION:
;           TELEPHONE: 312/474-6300
;           TELEFAX: 312/474-0448
;           TELEX: 25-3856
;   INFORMATION FOR SEQ ID NO: 206:
;       SEQUENCE CHARACTERISTICS:
;           LENGTH: 11 amino acids
;           TYPE: amino acid
;           TOPOLOGY: linear
;       MOLECULE TYPE: peptide
;       FEATURE:
;           NAME/KEY:  misc_feature
;           OTHER INFORMATION:  "XMP.350"
;       FEATURE:
;           NAME/KEY:  Modified-site
;           LOCATION:  C-Terminus
;           OTHER INFORMATION:  /label= Amidation
```

```
;           /note= "The C-Terminus is Amidated."
;       SEQUENCE DESCRIPTION: SEQ ID NO: 206:
US-09-765-527-206

  Query Match             100.0%;  Score 48;  DB 9;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 0.14;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          2 KKWLIQLKK 10


RESULT 12
US-09-881-490-181
; Sequence 181, Application US/09881490
; Patent No. US20020077298A1
;   GENERAL INFORMATION:
;       APPLICANT: Little II, Roger G.
;                  Lim, Edward
;                  Fadem, Mitchell B.
;       TITLE OF INVENTION: Anti-Fungal Peptides
;       NUMBER OF SEQUENCES: 211
;       CORRESPONDENCE ADDRESS:
;           ADDRESSEE: McAndrews, Held & Malloy, Ltd.
;           STREET: 500 West Madison Street, 34th FloorDrive
;           CITY: Chicago
;           STATE: Illinois
;           COUNTRY: United States of America
;           ZIP: 60661
;       COMPUTER READABLE FORM:
;           MEDIUM TYPE: Floppy disk
;           COMPUTER: IBM PC compatible
;           OPERATING SYSTEM: PC-DOS/MS-DOS
;           SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;           APPLICATION NUMBER: US/09/881,490
;           FILING DATE: 14-Jun-2001
;       PRIOR APPLICATION DATA:
;           APPLICATION NUMBER: 09/119,858
;           FILING DATE: <Unknown>
;           APPLICATION NUMBER: 08/372,105
;           FILING DATE: 13-JAN-95
;           APPLICATION NUMBER: 08/306,473
;           FILING DATE: 15-SEP-94
;           APPLICATION NUMBER: 08/273,540
;           FILING DATE: 11-JUL-94
;           APPLICATION NUMBER: 08/209,762
;           FILING DATE: 11-MAR-94
;           APPLICATION NUMBER: 08/183,222
;           FILING DATE: 14-JAN-94
;           APPLICATION NUMBER: 08/093,202
;           FILING DATE: 15-JUL-93
;           APPLICATION NUMBER: 08/030,644
;           FILING DATE: 12-MAR-93
;       ATTORNEY/AGENT INFORMATION:
;           NAME: McNicholas, Janet M.
;           REGISTRATION NUMBER: 32,918
;           REFERENCE/DOCKET NUMBER: 100-238/11021US01
;       TELECOMMUNICATION INFORMATION:
;           TELEPHONE: 312/707-8889
;           TELEFAX: 312/707-9155
;           TELEX: 650 388-1248
;   INFORMATION FOR SEQ ID NO: 181:
;       SEQUENCE CHARACTERISTICS:
;           LENGTH: 11 amino acids
;           TYPE: amino acid
;           TOPOLOGY: linear
;       MOLECULE TYPE: peptide
;       FEATURE:
;           NAME/KEY:  misc_feature
;           OTHER INFORMATION:  "XMP.350"
;       FEATURE:
;           NAME/KEY:  Modified-site
;           LOCATION:  C-Terminus
;           OTHER INFORMATION:  /label= Amidation
;           /note= "The C-Terminus is Amidated"
;       SEQUENCE DESCRIPTION: SEQ ID NO: 181:
US-09-881-490-181

  Query Match             100.0%;  Score 48;  DB 9;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 0.14;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||||||
Db          2 KKWLIQLKK 10


RESULT 13
US-09-765-527-192
; Sequence 192, Application US/09765527
; Patent No. US20020006638A1
;   GENERAL INFORMATION:
;       APPLICANT: Better, Marc D.
;       TITLE OF INVENTION: Methods for Recombinant Microbial Production of
;                           Fusion Proteins and BPI-Derived Peptides
;       NUMBER OF SEQUENCES: 265
;       CORRESPONDENCE ADDRESS:
;           ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
;           STREET: 6300 Sears Tower, 233 South Wacker Drive
;           CITY: Chicago
;           STATE: Illinois
;           COUNTRY: United States of America
;           ZIP: 60606-6402
;       COMPUTER READABLE FORM:
;           MEDIUM TYPE: Floppy disk
;           COMPUTER: IBM PC compatible
;           OPERATING SYSTEM: PC-DOS/MS-DOS
;           SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;           APPLICATION NUMBER: US/09/765,527
;           FILING DATE: 18-Jan-2001
;       PRIOR APPLICATION DATA:
;           APPLICATION NUMBER: 08/621,803
;           FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;           NAME: Borun, Michael F.
;           REGISTRATION NUMBER: 25,447
;           REFERENCE/DOCKET NUMBER: 27129/33199
;       TELECOMMUNICATION INFORMATION:
;           TELEPHONE: 312/474-6300
;           TELEFAX: 312/474-0448
;           TELEX: 25-3856
;   INFORMATION FOR SEQ ID NO: 192:
;       SEQUENCE CHARACTERISTICS:
;           LENGTH: 8 amino acids
;           TYPE: amino acid
;           TOPOLOGY: linear
;       MOLECULE TYPE: peptide
;       FEATURE:
;           NAME/KEY:  misc_feature
;           OTHER INFORMATION:  "XMP.326"
;       FEATURE:
;           NAME/KEY:  Modified-site
;           LOCATION:  C-Terminus
;           OTHER INFORMATION:  /label= Amidation
;           /note= "The C-Terminus is Amidated."
;       SEQUENCE DESCRIPTION: SEQ ID NO: 192:
US-09-765-527-192

  Query Match              89.6%;  Score 43;  DB 9;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 5.4e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLK 8
```

```
              ||||||||
Db          1 KKWLIQLK 8


RESULT 14
US-09-881-490-157
; Sequence 157, Application US/09881490
; Patent No. US20020077298A1
;   GENERAL INFORMATION:
;        APPLICANT: Little II, Roger G.
;                   Lim, Edward
;                   Fadem, Mitchell B.
;        TITLE OF INVENTION: Anti-Fungal Peptides
;        NUMBER OF SEQUENCES: 211
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: McAndrews, Held & Malloy, Ltd.
;             STREET: 500 West Madison Street, 34th FloorDrive
;             CITY: Chicago
;             STATE: Illinois
;             COUNTRY: United States of America
;             ZIP: 60661
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/881,490
;             FILING DATE: 14-Jun-2001
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/119,858
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: 08/372,105
;             FILING DATE: 13-JAN-95
;             APPLICATION NUMBER: 08/306,473
;             FILING DATE: 15-SEP-94
;             APPLICATION NUMBER: 08/273,540
;             FILING DATE: 11-JUL-94
;             APPLICATION NUMBER: 08/209,762
;             FILING DATE: 11-MAR-94
;             APPLICATION NUMBER: 08/183,222
;             FILING DATE: 14-JAN-94
;             APPLICATION NUMBER: 08/093,202
;             FILING DATE: 15-JUL-93
;             APPLICATION NUMBER: 08/030,644
;             FILING DATE: 12-MAR-93
;        ATTORNEY/AGENT INFORMATION:
;             NAME: McNicholas, Janet M.
;             REGISTRATION NUMBER: 32,918
;             REFERENCE/DOCKET NUMBER: 100-238/11021US01
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 312/707-8889
;             TELEFAX: 312/707-9155
;             TELEX: 650 388-1248
;   INFORMATION FOR SEQ ID NO: 157:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 8 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        FEATURE:
;             NAME/KEY:  misc_feature
;             OTHER INFORMATION:  "XMP.324"
;        FEATURE:
;             NAME/KEY:  Modified-site
;             LOCATION:  C-Terminus
;             OTHER INFORMATION:  /label= Amidation
;             /note= "The C-Terminus is Amidated"
;        SEQUENCE DESCRIPTION: SEQ ID NO: 157:
US-09-881-490-157

  Query Match             89.6%;  Score 43;  DB 9;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 5.4e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 KWLIQLKK 9
              ||||||||
Db          1 KWLIQLKK 8


RESULT 15
US-09-881-490-159
; Sequence 159, Application US/09881490
; Patent No. US20020077298A1
;   GENERAL INFORMATION:
;        APPLICANT: Little II, Roger G.
;                   Lim, Edward
;                   Fadem, Mitchell B.
;        TITLE OF INVENTION: Anti-Fungal Peptides
;        NUMBER OF SEQUENCES: 211
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: McAndrews, Held & Malloy, Ltd.
;             STREET: 500 West Madison Street, 34th FloorDrive
;             CITY: Chicago
;             STATE: Illinois
;             COUNTRY: United States of America
;             ZIP: 60661
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/881,490
;             FILING DATE: 14-Jun-2001
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/119,858
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: 08/372,105
;             FILING DATE: 13-JAN-95
;             APPLICATION NUMBER: 08/306,473
;             FILING DATE: 15-SEP-94
;             APPLICATION NUMBER: 08/273,540
;             FILING DATE: 11-JUL-94
;             APPLICATION NUMBER: 08/209,762
;             FILING DATE: 11-MAR-94
;             APPLICATION NUMBER: 08/183,222
;             FILING DATE: 14-JAN-94
;             APPLICATION NUMBER: 08/093,202
;             FILING DATE: 15-JUL-93
;             APPLICATION NUMBER: 08/030,644
;             FILING DATE: 12-MAR-93
;        ATTORNEY/AGENT INFORMATION:
;             NAME: McNicholas, Janet M.
;             REGISTRATION NUMBER: 32,918
;             REFERENCE/DOCKET NUMBER: 100-238/11021US01
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 312/707-8889
;             TELEFAX: 312/707-9155
;             TELEX: 650 388-1248
;   INFORMATION FOR SEQ ID NO: 159:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 8 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        FEATURE:
;             NAME/KEY:  misc_feature
;             OTHER INFORMATION:  "XMP.326"
;        FEATURE:
;             NAME/KEY:  Modified-site
;             LOCATION:  C-Terminus
;             OTHER INFORMATION:  /label= Amidation
;             /note= "The C-Terminus is Amidated"
;        SEQUENCE DESCRIPTION: SEQ ID NO: 159:
US-09-881-490-159
```

Query Match 89.6%; Score 43; DB 9; Length 8;
Best Local Similarity 100.0%; Pred. No. 5.4e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 KKWLIQLK 8
              ||||||||
Db          1 KKWLIQLK 8
```

Search completed: October 14, 2003, 14:44:51
Job time : 21.9375 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 14:33:18 ; Search time 11.8125 Seconds
(without alignments)
73.271 Million cell updates/sec

Title: US-09-881-490-164
Perfect score: 48
Sequence: 1 KKWLIQLKK 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	37	77.1	189	2	AH1029	probable membrane
2	37	77.1	212	2	T31058	hypothetical prote
3	37	77.1	275	2	H81074	phytoene synthase-
4	36	75.0	70	2	G64500	hypothetical prote
5	36	75.0	89	2	F91003	Kil protein [impor
6	36	75.0	578	2	B71431	hypothetical prote
7	35	72.9	198	2	T51962	Rac-like GTP bindi
8	35	72.9	585	2	C70341	acetolactate synth
9	35	72.9	731	2	A99106	hypothetical prote
10	35	72.9	2655	2	D96595	probable acetyl-Co
11	34	70.8	59	2	AE2023	hypothetical prote
12	34	70.8	89	2	B85638	probable Kil prote
13	34	70.8	89	2	C85848	hypothetical prote
14	34	70.8	180	2	T37857	hypothetical prote
15	34	70.8	210	2	B40203	4-alpha-glucanotra
16	34	70.8	276	2	D81868	probable transfera
17	34	70.8	285	2	A69748	hypothetical prote
18	34	70.8	287	2	B90540	conserved hypothet
19	34	70.8	382	2	S55166	hypothetical prote
20	34	70.8	547	2	A64991	hypothetical ABC t
21	34	70.8	547	2	D91016	hypothetical prote
22	34	70.8	547	2	F85860	hypothetical prote
23	34	70.8	4588	2	T28667	dynein beta heavy
24	34	70.8	4639	1	A54794	dynein heavy chain
25	34	70.8	4725	1	A44357	dynein heavy chain
26	33	68.8	53	2	S72388	pd53 protein - Ent
27	33	68.8	95	2	I39575	nickel-cobalt-cadm
28	33	68.8	185	2	D82418	MutT/nudix family
29	33	68.8	185	2	A96989	transcription regu
30	33	68.8	201	2	T48862	rac-like protein A
31	33	68.8	209	2	T05164	hypothetical prote
32	33	68.8	220	2	AB0084	probable DedA-fami
33	33	68.8	233	2	D96763	hypothetical prote
34	33	68.8	252	2	F97072	ABC transporter, A
35	33	68.8	264	2	AE2787	exopolysaccharide
36	33	68.8	264	2	G97566	exopolysaccharide
37	33	68.8	335	2	A89834	teichoic acid bios
38	33	68.8	344	2	S46871	B13R protein - var
39	33	68.8	344	2	E72173	D2R protein - vari
40	33	68.8	344	2	T28611	hypothetical prote
41	33	68.8	367	2	AD3631	nitrous-oxide redu
42	33	68.8	396	2	T46901	hypothetical prote
43	33	68.8	585	2	A39286	parathyroid hormon
44	33	68.8	589	2	I59297	parathyroid hormon
45	33	68.8	591	2	S44203	parathyroid hormon

ALIGNMENTS

RESULT 1
AH1029
probable membrane protein STY4561 [imported] - Salmonella enterica subsp. enterica
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AH1029
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churth, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Fa, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AH1029
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-189 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD09337.1; PID:g16505337; GSPDB:GN00176
C;Genetics:
A;Gene: STY4561

Query Match 77.1%; Score 37; DB 2; Length 189;
Best Local Similarity 66.7%; Pred. No. 13;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 KKWLIQLKK 9
              |:|||| |:
Db        101 KRWLIQAKR 109
```

RESULT 2
T31058
hypothetical protein C05D9.5 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 17-Mar-2000
C;Accession: T31058
R;Fulton, L.; Gattung, S.
submitted to the EMBL Data Library, September 1999
A;Description: The sequence of C. elegans cosmid C05D9.
A;Reference number: Z20960
A;Accession: T31058
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-212 <FUL>
A;Cross-references: EMBL:U64605; PIDN:AAB04584.1
A;Experimental source: strain Bristol N2
C;Genetics:
A;Map position: X
A;Introns: 21/3; 82/1; 161/2
A;Note: C05D9.5
C;Superfamily: translation initiation factor eIF-4E

```
  Query Match             77.1%;  Score 37;  DB 2;  Length 212;
  Best Local Similarity   75.0%;  Pred. No. 14;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 KWLIQLKK 9
              ||:|:|||
Db        110 KWIIRLKK 117


RESULT 3
H81074
phytoene synthase-related protein NMB1521 [imported] - Neisseria meningitidis (strain MC
C;Species: Neisseria meningitidis
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C;Accession: H81074
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A;Reference number: A81000; MUID:20175755; PMID:10710307
A;Accession: H81074
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-275 <TET>
A;Cross-references: GB:AE002501; GB:AE002098; NID:g7226755; PIDN:AAF41877.1; PID:g722676
A;Experimental source: serogroup B, strain MC58
C;Genetics:
A;Gene: NMB1521

  Query Match             77.1%;  Score 37;  DB 2;  Length 275;
  Best Local Similarity   77.8%;  Pred. No. 19;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              | ||| |||
Db        263 KDWLIMLKK 271


RESULT 4
G64500
hypothetical protein MJ1608 - Methanococcus jannaschii
C;Species: Methanococcus jannaschii
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
C;Accession: G64500
R;Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,
; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A;Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C
A;Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
A;Reference number: A64300; MUID:96337999; PMID:8688087
A;Accession: G64500
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-70 <BUL>
A;Cross-references: GB:U67601; GB:L77117; NID:g2826439; PIDN:AAB99635.1; PID:g1500506; T
C;Genetics:
A;Map position: FOR1582918-1583130

  Query Match             75.0%;  Score 36;  DB 2;  Length 70;
  Best Local Similarity   75.0%;  Pred. No. 7.4;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLK 8
              |||:|:||
Db         63 KKWVIKLK 70


RESULT 5
F91003
Kil protein [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509952)
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 17-May-2002
C;Accession: F91003
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han,
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: F91003
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-89 <HAY>
A;Cross-references: GB:BA000007; PIDN:BAB36421.1; PID:g13362467; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs2998
C;Superfamily: phage lambda kil protein

  Query Match             75.0%;  Score 36;  DB 2;  Length 89;
  Best Local Similarity   55.6%;  Pred. No. 9.3;
  Matches    5;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||:: ||:
Db         74 KKWILMLKR 82


RESULT 6
B71431
hypothetical protein - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
A;Variety: columbia
C;Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 05-Dec-1998
C;Accession: B71431
R;Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.;
 P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenegger, T.; Pohl, T.M.; Terryn, N.;
avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk,
Nature 391, 485-488, 1998
A;Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdome
erhoft, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.;
C.; Chalwatzis, N.
A;Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis
A;Reference number: A71400; MUID:98121113; PMID:9461215
A;Accession: B71431
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-578 <BEV>
A;Cross-references: GB:Z97341; NID:g2244991; PID:e326992; PID:g2245000
C;Genetics:
A;Map position: 4COP9-4G3845

  Query Match             75.0%;  Score 36;  DB 2;  Length 578;
  Best Local Similarity   87.5%;  Pred. No. 59;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 KWLIQLKK 9
              ||| ||||
Db         30 KWLDQLKK 37


RESULT 7
T51962
Rac-like GTP binding protein [imported] - Picea mariana
C;Species: Picea mariana
C;Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 03-Nov-2000
C;Accession: T51962
R;Perry, D.J.; Bousquet, J.
Genetics 149, 1089-1098, 1998
A;Title: Sequence-tagged-site (STS) markers of arbitrary genes. Development, charact
A;Reference number: Z25268; MUID:98278823; PMID:9611216
A;Accession: T51962
A;Status: preliminary; translated from GB/EMBL/DDBJ
```

```
A;Molecule type: mRNA
A;Residues: 1-198 <PER>
A;Cross-references: EMBL:AF051223; PIDN:AAC32124.1
C;Genetics:
A;Note: Sb30
C;Superfamily: ras transforming protein; translation elongation factor Tu homology

  Query Match             72.9%;  Score 35;  DB 2;  Length 198;
  Best Local Similarity   62.5%;  Pred. No. 32;
  Matches    5;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLK 8
              ||||::|:
Db         98 KKWLLELR 105


RESULT 8
C70341
acetolactate synthase (EC 4.1.3.18) large chain - Aquifex aeolicus
C;Species: Aquifex aeolicus
C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 15-Oct-1999
C;Accession: C70341
R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov
V.
Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A;Reference number: A70300; MUID:98196666; PMID:9537320
A;Accession: C70341
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-585 <AQF>
A;Cross-references: GB:AE000690; NID:g2983100; PIDN:AAC06706.1; PID:g2983101; GB:AE00065
A;Experimental source: strain VF5
C;Genetics:
A;Gene: ilvB
C;Superfamily: acetolactate synthase large chain; thiamin pyrophosphate-binding domain h
C;Keywords: branched-chain amino acid biosynthesis; carbon-carbon lyase; flavoprotein; m
F;440-488/Domain: thiamin pyrophosphate-binding domain homology <TPB>

  Query Match             72.9%;  Score 35;  DB 2;  Length 585;
  Best Local Similarity   66.7%;  Pred. No. 93;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||| |::|
Db        351 KKWLEQIEK 359


RESULT 9
A99106
hypothetical protein orf731 [imported] - Guillardia theta nucleomorph
C;Species: nucleomorph Guillardia theta
A;Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001
C;Accession: A99106
R;Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Rei
Nature 410, 1091-1096, 2001
A;Title: The highly reduced genome of an enslaved algal nucleus.
A;Reference number: A99082; MUID:11323671; PMID:11323671
A;Accession: A99106
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-731 <DOU>
A;Cross-references: GB:AJ010592; NID:g12580677; PIDN:CAC26995.1; GSPDB:GN00151
C;Genetics:
A;Gene: orf731
A;Map position: 2
A;Genome: nucleomorph
C;Keywords: nucleomorph

  Query Match             72.9%;  Score 35;  DB 2;  Length 731;
  Best Local Similarity   66.7%;  Pred. No. 1.2e+02;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||:  |||
Db        273 KKWIFLLKK 281


RESULT 10
D96595
probable acetyl-CoA synthetase, 45051-31547 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: D96595
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alc
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kir
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marz:
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tal
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: D96595
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-2655 <STO>
A;Cross-references: GB:AE005173; NID:g10645429; PIDN:AAG21546.1; GSPDB:GN00141
C;Genetics:
A;Gene: F7A10.14
A;Map position: 1

  Query Match             72.9%;  Score 35;  DB 2;  Length 2655;
  Best Local Similarity   66.7%;  Pred. No. 4.1e+02;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              ||| |||::
Db       2335 KKWPIQLRR 2343


RESULT 11
AE2023
hypothetical protein asl1739 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C;Accession: AE2023
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iri
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Taba
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacteriu
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AE2023
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-59 <KUR>
A;Cross-references: GB:BA000019; PIDN:BAB73438.1; PID:g17130829; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: asl1739

  Query Match             70.8%;  Score 34;  DB 2;  Length 59;
  Best Local Similarity   55.6%;  Pred. No. 15;
  Matches    5;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              | ||::||:
Db         47 KAWLVKLKR 55


RESULT 12
B85638
```

probable Kil protein of bacteriphage BP-933W kilW [imported] - Escherichia coli (strain
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 17-May-2002
C;Accession: B85638
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: B85638
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-89 <STO>
A;Cross-references: GB:AE005174; NID:g12514292; PIDN:AAG55566.1; GSPDB:GN00145; UWGP:Z14
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: kilW
C;Superfamily: phage lambda kil protein

Query Match 70.8%; Score 34; DB 2; Length 89;
Best Local Similarity 62.5%; Pred. No. 22;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KKWLIQLK 8
||||:: ||
Db 74 KKWILMLK 81

RESULT 13
C85848
hypothetical protein Z3364 [imported] - Escherichia coli (strain O157:H7, substrain EDL9
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 17-May-2002
C;Accession: C85848
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: C85848
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-89 <STO>
A;Cross-references: GB:AE005174; NID:g12516417; PIDN:AAG57247.1; GSPDB:GN00145; UWGP:Z33
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: Z3364
C;Superfamily: phage lambda kil protein

Query Match 70.8%; Score 34; DB 2; Length 89;
Best Local Similarity 62.5%; Pred. No. 22;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KKWLIQLK 8
||||:: ||
Db 74 KKWILMLK 81

RESULT 14
T37857
hypothetical protein SPAC17G8.05 - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C;Accession: T37857
R;Devlin, K.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, February 1996
A;Reference number: Z21750
A;Accession: T37857
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-180 <DEV>
A;Cross-references: EMBL:Z69795; PIDN:CAA93688.1; GSPDB:GN00068; SPDB:SPAC17G8.05
A;Experimental source: strain 972h-; cosmid c17G8
C;Genetics:
A;Gene: SPDB:SPAC17G8.05
A;Map position: 1
A;Introns: 22/1

Query Match 70.8%; Score 34; DB 2; Length 180;
Best Local Similarity 62.5%; Pred. No. 44;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KKWLIQLK 8
||||::| |
Db 22 KKWVVQYK 29

RESULT 15
B40203
4-alpha-glucanotransferase (EC 2.4.1.25) / amylo-1,6-glucosidase (EC 3.2.1.33) - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 28-Aug-1992 #sequence_revision 28-Aug-1992 #text_change 16-Aug-2002
C;Accession: B40203
R;Yang, B.Z.; Ding, J.H.; Enghild, J.J.; Bao, Y.; Chen, Y.T.
J. Biol. Chem. 267, 9294-9299, 1992
A;Title: Molecular cloning and nucleotide sequence of cDNA encoding human muscle gly
A;Reference number: A40203; MUID:92250533; PMID:1374391
A;Accession: B40203
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-210 <YAN>
C;Superfamily: 4-alpha-glucanotransferase / amylo-1,6-glucosidase
C;Keywords: glycosidase; glycosyltransferase; hexosyltransferase; hydrolase; polysac

Query Match 70.8%; Score 34; DB 2; Length 210;
Best Local Similarity 62.5%; Pred. No. 52;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 KWLIQLKK 9
|||::|:|
Db 203 KWLLELQK 210

Search completed: October 14, 2003, 14:38:06
Job time : 11.8125 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 14:33:18 ; Search time 6.375 Seconds
(without alignments)
66.391 Million cell updates/sec

Title: US-09-881-490-164
Perfect score: 48
Sequence: 1 KKWLIQLKK 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	37	77.1	212	1	IFE4_CAEEL	Q22888 caenorhabdi
2	36	75.0	70	1	YG08_METJA	Q59003 methanococc
3	34	70.8	180	1	YD55_SCHPO	Q10317 schizosacch
4	34	70.8	245	1	IFE3_HUMAN	O60573 h eukaryoti
5	34	70.8	245	1	IFE3_MOUSE	Q8bmb3 mus musculu
6	34	70.8	382	1	YJO7_YEAST	P47007 saccharomyc
7	34	70.8	430	1	SECY_STAAU	O08387 staphylococ
8	34	70.8	547	1	YCJ1_ECOLI	P33941 escherichia
9	34	70.8	4639	1	DYHC_DROME	P37276 drosophila
10	34	70.8	4725	1	DYHC_DICDI	P34036 dictyostel
11	33	68.8	95	1	NCCY_ALCXX	Q44581 alcaligenes
12	33	68.8	201	1	RAC2_ARATH	Q38903 arabidopsis
13	33	68.8	209	1	V282_ARATH	O65421 arabidopsis
14	33	68.8	218	1	GCH1_PASMU	P57865 pasteurella
15	33	68.8	267	1	EXOR_RHILV	Q52822 rhizobium l
16	33	68.8	344	1	SPI2_VARV	P33830 variola vir
17	33	68.8	396	1	PTD4_HUMAN	Q9ntk5 homo sapien
18	33	68.8	396	1	PTD4_MOUSE	Q9cz30 mus musculu
19	33	68.8	585	1	PTRR_DIDMA	P25107 didelphis m
20	33	68.8	585	1	PTRR_PIG	P50133 sus scrofa
21	33	68.8	591	1	PTRR_MOUSE	P41593 mus musculu
22	33	68.8	591	1	PTRR_RAT	P25961 rattus norv
23	33	68.8	593	1	PTRR_HUMAN	Q03431 homo sapien
24	33	68.8	639	1	NOSZ_RHIME	Q59746 rhizobium m
25	33	68.8	650	1	MI10_CAEEL	P34400 caenorhabdi
26	33	68.8	1742	1	MY5C_HUMAN	Q9nqx4 homo sapien
27	33	68.8	1877	1	PCK5_MOUSE	Q04592 mus musculu
28	33	68.8	2109	1	RRPL_VSVJH	P13615 vesicular s
29	32	66.7	290	1	Y084_MYCGE	P47330 mycoplasma
30	32	66.7	387	1	BODG_MOUSE	Q924y0 mus musculu
31	32	66.7	387	1	BODG_RAT	Q9qzu7 rattus norv
32	32	66.7	423	1	AATM_CHICK	P00508 gallus gall
33	32	66.7	454	1	Y095_MYCPN	P75538 mycoplasma
34	32	66.7	470	1	MRS2_YEAST	Q01926 saccharomyc
35	32	66.7	538	1	TH1P_HAEIN	P44985 haemophilus
36	32	66.7	1067	1	CARB_CLOPE	Q8xhb3 clostridium
37	32	66.7	1328	1	POLX_TOBAC	P10978 nicotiana t
38	31.5	65.6	1953	1	BNI1_YEAST	P41832 saccharomyc
39	31	64.6	89	1	VKIL_LAMBD	P03758 bacteriopha
40	31	64.6	113	1	YB51_METJA	Q58551 methanococc
41	31	64.6	124	1	LBD9_ARATH	O82198 arabidopsis
42	31	64.6	152	1	FGF1_PIG	P20002 sus scrofa
43	31	64.6	155	1	FGF1_CHICK	P19596 gallus gall
44	31	64.6	155	1	FGF1_HUMAN	P05230 homo sapien
45	31	64.6	155	1	FGF1_MESAU	P34004 mesocricetu

ALIGNMENTS

```
RESULT 1
IFE4_CAEEL
ID   IFE4_CAEEL     STANDARD;      PRT;   212 AA.
AC   Q22888;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Eukaryotic translation initiation factor 4E (eIF-4E) (eIF4E) (mRNA
DE   cap-binding protein) (eIF-4F 25 kDa subunit).
GN   IFE-4 OR C05D9.5.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RX   MEDLINE=20209441; PubMed=10744754;
RA   Keiper B.D., Lamphear B.J., Deshpande A.M., Jankowska-Anyszka M.,
RA   Aamodt E.J., Blumenthal T., Rhoads R.E.;
RT   "Functional characterization of five eIF4E isoforms in Caenorhabditis
RT   elegans.";
RL   J. Biol. Chem. 275:10590-10596(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Fulton L., Gattung S.;
RL   Submitted (JUL-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: RECOGNIZES AND BINDS THE 7-METHYLGUANOSINE-CONTAINING
CC       MRNA "CAP" DURING AN EARLY STEP IN THE INITIATION OF PROTEIN
CC       SYNTHESIS AND FACILITATES RIBOSOME BINDING BY INDUCING THE
CC       UNWINDING OF THE MRNAS SECONDARY STRUCTURES.
CC   -!- SUBUNIT: EIF4F IS A TRIMER COMPOSED OF EIF4E, EIF4G AND EIF4A
CC       (WHICH CAN CYCLE IN AND OUT OF THE COMPLEX).
CC   -!- SIMILARITY: BELONGS TO THE EUKARYOTIC INITIATION FACTOR 4E FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF214651; AAF62414.1; -.
DR   EMBL; U64605; AAF98601.1; -.
DR   PIR; T31058; T31058.
DR   HSSP; P07260; 1AP8.
DR   WormPep; C05D9.5; CE07931.
DR   InterPro; IPR001040; TIF_eIF_4E.
DR   Pfam; PF01652; IF4E; 1.
DR   ProDom; PD003697; TIF_eIF_4E; 1.
DR   PROSITE; PS00813; IF4E; 1.
KW   Initiation factor; Protein biosynthesis; RNA-binding;
KW   Multigene family.
SQ   SEQUENCE   212 AA;  24584 MW;  87BA3DFE9FE8A486 CRC64;
```

```
  Query Match             77.1%;  Score 37;  DB 1;  Length 212;
  Best Local Similarity   75.0%;  Pred. No. 5.1;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 KWLIQLKK 9
              ||:|:|||
Db        110 KWIIRLKK 117


RESULT 2
YG08_METJA
ID   YG08_METJA     STANDARD;      PRT;    70 AA.
AC   Q59003;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical protein MJ1608.
GN   MJ1608.
OS   Methanococcus jannaschii.
OC   Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC   Methanocaldococcaceae; Methanocaldococcus.
OX   NCBI_TaxID=2190;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX   MEDLINE=96337999; PubMed=8688087;
RA   Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA   Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA   Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA   Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA   Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA   Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA   Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA   Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT   "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT   jannaschii.";
RL   Science 273:1058-1073(1996).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U67601; AAB99635.1; -.
DR   PIR; G64500; G64500.
DR   TIGR; MJ1608; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   70 AA;  8122 MW;  BDC6BE5A1A47AF57 CRC64;

  Query Match             75.0%;  Score 36;  DB 1;  Length 70;
  Best Local Similarity   75.0%;  Pred. No. 2.7;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLK 8
              |||:|:||
Db         63 KKWVIKLK 70


RESULT 3
YD65_SCHPO
ID   YD65_SCHPO     STANDARD;      PRT;    180 AA.
AC   Q10317;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hypothetical protein C17G8.05 in chromosome I.
GN   SPAC17G8.05.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RX   MEDLINE=21848401; PubMed=11859360;
RA   Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA   Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA   Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA   Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA   Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA   Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA   James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA   Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA   Oliver K., O'Neil S., Pearson D., Quail M.A., Rabbinowitsch E.,
RA   Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA   Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA   Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA   Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA   Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA   Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
RA   Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA   Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA   Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA   Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA   Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA   Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA   Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA   Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA   Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT   "The genome sequence of Schizosaccharomyces pombe.";
RL   Nature 415:871-880(2002).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z69795; CAA93688.1; -.
DR   PIR; T37857; T37857.
DR   GeneDB_SPombe; SPAC17G8.05; -.
KW   Hypothetical protein.
SQ   SEQUENCE   180 AA;  20931 MW;  A471D5243297E0B2 CRC64;

  Query Match             70.8%;  Score 34;  DB 1;  Length 180;
  Best Local Similarity   62.5%;  Pred. No. 16;
  Matches    5;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLK 8
              |||::| |
Db         22 KKWVVQYK 29


RESULT 4
IFE3_HUMAN
ID   IFE3_HUMAN     STANDARD;      PRT;    245 AA.
AC   O60573; O75349;
DT   15-SEP-2003 (Rel. 42, Created)
DT   15-SEP-2003 (Rel. 42, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Eukaryotic translation initiation factor 4E type 3 (eIF-4E type 3)
DE   (eIF4E type 3) (mRNA cap-binding protein type 3) (Eukaryotic
DE   translation initiation factor 4E-like 3) (Eukaryotic translation
DE   initiation factor 4E homologous protein) (mRNA cap-binding protein
DE   4EHP) (eIF4E-like protein 4E-LP).
GN   EIF4EL3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., CHARACTERIZATION, 3D-STRUCTURE MODELING, AND
RP   MUTAGENESIS.
RC   TISSUE=Follicle;
RX   MEDLINE=98250763; PubMed=9582349;
RA   Rom E., Kim H.C., Gingras A.-C., Marcotrigiano J., Favre D., Olsen H.,
RA   Burley S.K., Sonenberg N.;
RT   "Cloning and characterization of 4EHP, a novel mammalian eIF4E-related
RT   cap-binding protein.";
RL   J. Biol. Chem. 273:13104-13109(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Breast;
RA   Joshi B., Norris K.K., Jagus R.;
RT   "Isolation of human and mouse cDNAs encoding novel eIF4E-like
RT   proteins.";
RL   Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Blood;
RX   MEDLINE=98318631; PubMed=9653160;
RA   Mao M., Fu G., Wu J.-S., Zhang Q.-H., Zhou J., Kan L.-X., Huang Q.-H.,
RA   He K.-L., Gu B.-W., Han Z.-G., Shen Y., Gu J., Yu Y.-P., Xu S.-H.,
RA   Wang Y.-X., Chen S.-J., Chen Z.;
RT   "Identification of genes expressed in human CD34(+) hematopoietic
RT   stem/progenitor cells by expressed sequence tags and efficient full-
RT   length cDNA cloning.";
RL   Proc. Natl. Acad. Sci. U.S.A. 95:8175-8180(1998).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Muscle, Urinary bladder, and Uterus;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: Recognizes and binds the 7-methylguanosine-containing
CC       mRNA "CAP" during an early step in the initiation of protein
CC       synthesis and facilitates ribosome binding by inducing the
CC       unwinding of the mRNAS secondary structures.
CC   -!- SUBUNIT: EIF4F is a trimer composed of EIF4E, EIF4G and EIF4A
CC       (which can cycle in and out of the complex) (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE EUKARYOTIC INITIATION FACTOR 4E FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF047695; AAC18565.1; -.
DR   EMBL; AF068117; AAC19374.1; -.
DR   EMBL; AF038957; AAC39871.1; -.
DR   EMBL; BC005392; AAH05392.1; -.
DR   EMBL; BC005874; AAH05874.1; -.
DR   EMBL; BC021226; AAH21226.1; -.
DR   EMBL; BC021690; AAH21690.1; -.
DR   Genew; HGNC:3293; EIF4EL3.
DR   MIM; 605895; -.
DR   GO; GO:0003731; F:mRNA cap binding; TAS.
DR   GO; GO:0006445; P:regulation of translation; TAS.
DR   InterPro; IPR001040; TIF_eIF_4E.
DR   Pfam; PF01652; IF4E; 1.
DR   ProDom; PD003697; TIF_eIF_4E; 1.
DR   PROSITE; PS00813; IF4E; 1.
KW   Initiation factor; Protein biosynthesis; RNA-binding;
KW   Multigene family.
FT   MUTAGEN      63     63       W->A: UNABLE TO BIND CAPPED RNA.
FT   MUTAGEN      95     95       W->A: ABILITY TO BIND CAPPED MRNA REDUCED
FT                                TO 40% OF WILD-TYPE.
FT   MUTAGEN     124    126       WED->FAA: UNABLE TO BIND CAPPED RNA.
FT   MUTAGEN     124    124       W->A: ABILITY TO BIND CAPPED MRNA REDUCED
FT                                TO LESS THAN 10% OF WILD-TYPE.
FT   MUTAGEN     124    124       W->F: ABILITY TO BIND CAPPED MRNA REDUCED
FT                                TO 13% OF WILD-TYPE.
FT   MUTAGEN     125    125       E->A: ABILITY TO BIND CAPPED MRNA REDUCED
FT                                TO LESS THAN 10% OF WILD-TYPE.
FT   MUTAGEN     126    126       D->A: SLIGHT REDUCTION IN ABILITY TO BIND
FT                                CAPPED RNA.
FT   MUTAGEN     135    135       W->A: UNABLE TO BIND CAPPED RNA.
FT   MUTAGEN     148    148       W->A: UNABLE TO BIND CAPPED RNA.
FT   MUTAGEN     183    183       W->A: ABILITY TO BIND CAPPED MRNA REDUCED
FT                                TO LESS THAN 10% OF WILD-TYPE.
FT   MUTAGEN     183    183       W->F: UNABLE TO BIND CAPPED RNA.
FT   CONFLICT      1     27       MNNKFDALKDDDSGDHDQNEENSTQKD -> MMTVGTMIRM
FT                                KKTAHRKI (IN REF. 3).
SQ   SEQUENCE   245 AA;  28362 MW;  3D3075BFA48B3C12 CRC64;

  Query Match             70.8%;  Score 34;  DB 1;  Length 245;
  Best Local Similarity   62.5%;  Pred. No. 22;
  Matches    5;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 KWLIQLKK 9
              ||:|:|:|
Db        134 KWIIRLRK 141


RESULT 5
IFE3_MOUSE
ID   IFE3_MOUSE     STANDARD;      PRT;   245 AA.
AC   Q8BMB3; O88503;
DT   15-SEP-2003 (Rel. 42, Created)
DT   15-SEP-2003 (Rel. 42, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Eukaryotic translation initiation factor 4E type 3 (eIF-4E type 3)
DE   (eIF4E type 3) (mRNA cap-binding protein type 3) (Eukaryotic
DE   translation initiation factor 4E-like 3) (eIF4E-like protein 4E-LP).
GN   EIF4EL3.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NIH Swiss;
RA   Joshi B., Jagus R.;
RT   "Isolation of human and mouse cDNAs encoding novel eIF4E-like
RT   proteins.";
RL   Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Mesonephros;
RX   MEDLINE=22354683; PubMed=12466851;
RA   Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA   Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA   Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA   Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
```

```
RA   Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA   Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA   Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA   Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA   Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA   Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA   Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA   Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA   Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA   Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA   Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA   Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA   Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA   Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA   Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA   Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA   Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA   Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA   Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA   Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA   Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA   Birney E., Hayashizaki Y.;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
CC   -!- FUNCTION: Recognizes and binds the 7-methylguanosine-containing
CC       mRNA "CAP" during an early step in the initiation of protein
CC       synthesis and facilitates ribosome binding by inducing the
CC       unwinding of the mRNAS secondary structures (By similarity).
CC   -!- SUBUNIT: EIF4F is a trimer composed of EIF4E, EIF4G and EIF4A
CC       (which can cycle in and out of the complex) (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE EUKARYOTIC INITIATION FACTOR 4E FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF068116; AAC19373.1; -.
DR   EMBL; AK032965; BAC28102.1; -.
DR   MGD; MGI:1914440; Eif4el3.
DR   InterPro; IPR001040; TIF_eIF_4E.
DR   Pfam; PF01652; IF4E; 1.
DR   ProDom; PD003697; TIF_eIF_4E; 1.
DR   PROSITE; PS00813; 1F4E; 1.
KW   Initiation factor; Protein biosynthesis; RNA-binding;
KW   Multigene family.
FT   CONFLICT    200    200       L -> F (IN REF. 1).
SQ   SEQUENCE   245 AA;  28263 MW;  AA722843721A01CC CRC64;

  Query Match             70.8%;  Score 34;  DB 1;  Length 245;
  Best Local Similarity   62.5%;  Pred. No. 22;
  Matches    5;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 KWLIQLKK 9
              |:|:|:|
Db        134 KWIIRLRK 141


RESULT 6
YJO7_YEAST
ID   YJO7_YEAST     STANDARD;      PRT;   382 AA.
AC   P47007;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   Hypothetical 44.9 kDa protein in INO1-IDS2 intergenic region.
GN   YJL147C OR J0639.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c / FY1679;
RX   MEDLINE=96408771; PubMed=8813765;
RA   Katsoulou C., Tzermia M., Tavernarakis N., Alexandraki D.;
RT   "Sequence analysis of a 40.7 kb segment from the left arm of yeast
RT   chromosome X reveals 14 known genes and 13 new open reading frames
RT   including homologues of genes clustered on the right arm of
RT   chromosome XI.";
RL   Yeast 12:787-797(1996).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z49422; CAA89442.1; -.
DR   EMBL; X87371; CAA60808.1; -.
DR   PIR; S55166; S55166.
DR   SGD; S0003683; YJL147C.
KW   Hypothetical protein.
SQ   SEQUENCE   382 AA;  44862 MW;  DDAAF588AF9A3234 CRC64;

  Query Match             70.8%;  Score 34;  DB 1;  Length 382;
  Best Local Similarity   66.7%;  Pred. No. 34;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||||  : :
Db        298 KKWLIDIVK 306


RESULT 7
SECY_STAAU
ID   SECY_STAAU     STANDARD;      PRT;   430 AA.
AC   O08387;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Preprotein translocase secY subunit.
GN   SECY.
OS   Staphylococcus aureus.
OC   Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX   NCBI_TaxID=1280;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NCTC 8325;
RA   Segarra R.A., Iandolo J.J.;
RL   Submitted (APR-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: INVOLVED IN PROTEIN EXPORT. INTERACTS WITH SECA AND SECE
CC       TO ALLOW THE TRANSLOCATION OF PROTEINS ACROSS THE PLASMA MEMBRANE,
CC       BY FORMING PART OF A CHANNEL.
CC   -!- SUBUNIT: ONE OF SEVEN SECRETORY PROTEINS (SECA-F & SECY) THAT
CC       COMPRISE THE PROKARYOTIC PROTEIN TRANSLOCATION APPARATUS.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO THE SECY/SEC61-ALPHA FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U96620; AAB54022.1; -.
DR   InterPro; IPR002208; SecY.
DR   Pfam; PF00344; secY; 1.
```

```
DR   PRINTS; PR00303; SECYTRNLCASE.
DR   TIGRFAMs; TIGR00967; 3a0501s007; 1.
DR   PROSITE; PS00755; SECY_1; 1.
DR   PROSITE; PS00756; SECY_2; 1.
KW   Protein transport; Translocation; Transmembrane.
SQ   SEQUENCE   430 AA;  47612 MW;  AF6923FE32434031 CRC64;

  Query Match             70.8%;  Score 34;  DB 1;  Length 430;
  Best Local Similarity   75.0%;  Pred. No. 38;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLK 8
              ||||| |:
Db        333 KKWLITLR 340


RESULT 8
YOJI_ECOLI
ID   YOJI_ECOLI     STANDARD;      PRT;   547 AA.
AC   P33941; P33942;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hypothetical ABC transporter ATP-binding protein yojI.
GN   YOJI OR B2211.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / BHB2600;
RA   Richterich P., Lakey N., Gryan G., Jaehn L., Mintz L., Robison K.,
RA   Church G.M.;
RL   Submitted (OCT-1993) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=97251358; PubMed=9097040;
RA   Itoh T., Aiba H., Baba T., Fujita K., Hayashi K., Inada T.,
RA   Isono K., Kasai H., Kimura S., Kitakawa M., Kitagawa M.,
RA   Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA   Nakade S., Nakamura Y., Nashimoto H., Nishio Y., Oshima T.,
RA   Saito N., Sampei G., Seki Y., Sivasundaram S., Tagami H.,
RA   Takeda J., Takemoto K., Wada C., Yamamoto Y., Horiuchi T.;
RT   "A 460-kb DNA sequence of the Escherichia coli K-12 genome
RT   corresponding to the 40.1-50.0 min region on the linkage map.";
RL   DNA Res. 3:379-392(1996).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC       (Probable).
CC   -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. SIMILAR TO
CC       P.SYRINGAE SYRD.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U00008; AAA16403.1; -.
DR   EMBL; AE000310; AAC75271.1; -.
DR   EMBL; D90850; BAA15994.1; -.
DR   PIR; A64991; A64991.
DR   EcoGene; EG12070; yojI.
DR   InterPro; IPR003593; AAA_ATPase.
DR   InterPro; IPR001140; ABC_TM_transpt.
DR   InterPro; IPR003439; ABC_transporter.
DR   InterPro; IPR005898; Cyc_pep_transpt.
DR   Pfam; PF00664; ABC_membrane; 1.
DR   Pfam; PF00005; ABC_tran; 1.
DR   ProDom; PD000006; ABC_transporter; 1.
DR   SMART; SM00382; AAA; 1.
DR   TIGRFAMs; TIGR01194; cyc_pep_trnsptr; 1.
DR   PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR   PROSITE; PS50893; ABC_TRANSPORTER_2; 1.
KW   Hypothetical protein; ATP-binding; Transport; Transmembrane;
KW   Inner membrane; Complete proteome.
FT   TRANSMEM     16     36       POTENTIAL.
FT   TRANSMEM     48     68       POTENTIAL.
FT   TRANSMEM    125    145       POTENTIAL.
FT   TRANSMEM    149    169       POTENTIAL.
FT   TRANSMEM    240    260       POTENTIAL.
FT   TRANSMEM    270    290       POTENTIAL.
FT   NP_BIND     356    363       ATP (POTENTIAL).
SQ   SEQUENCE   547 AA;  61552 MW;  11523935EDFD9EAB CRC64;

  Query Match             70.8%;  Score 34;  DB 1;  Length 547;
  Best Local Similarity   75.0%;  Pred. No. 48;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLK 8
              :||| |||
Db        426 EKWLAQLK 433


RESULT 9
DYHC_DROME
ID   DYHC_DROME     STANDARD;      PRT;  4639 AA.
AC   P37276;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Dynein heavy chain, cytosolic (DYHC).
GN   CDHC OR DHC64C.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94375524; PubMed=8089180;
RA   Li M., McGrail M., Serr M., Hays T.S.;
RT   "Drosophila cytoplasmic dynein, a microtubule motor that is
RT   asymmetrically localized in the oocyte.";
RL   J. Cell Biol. 126:1475-1494(1994).
RN   [2]
RP   SEQUENCE OF 1877-1998 FROM N.A.
RX   MEDLINE=94243034; PubMed=8186464;
RA   Rasmusson K., Serr M., Gepner J., Gibbons I., Hays T.S.;
RT   "A family of dynein genes in Drosophila melanogaster.";
RL   Mol. Biol. Cell 5:45-55(1994).
CC   -!- FUNCTION: Cytoplasmic dynein acts as a motor for the intracellular
CC       retrograde motility of vesicles and organelles along microtubules.
CC       Dynein has ATPase activity; the force-producing power stroke is
CC       thought to occur on release of ADP.
CC   -!- SUBUNIT: Consists of at least two heavy chains and a number of
CC       intermediate and light chains.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- DOMAIN: Dynein heavy chains probably consist of an N-terminal stem
CC       (which binds cargo and interacts with other dynein components),
CC       and the head or motor domain. The motor contains six tandemly-
CC       linked AAA domains in the head, which form a ring. A stalk-like
CC       structure (formed by two of the coiled coil domains) protrudes
```

```
CC       between AAA 4 and AAA 5 and terminates in a microtubule-binding
CC       site. A seventh domain may also contribute to this ring; it is not
CC       clear whether the N-terminus or the C-terminus forms this extra
CC       domain. There are four well-conserved and two non-conserved ATPase
CC       sites, one per AAA domain. Probably only one of these (within AAA
CC       1) actually hydrolyzes ATP, the others may serve a regulatory
CC       function.
CC   -!- SIMILARITY: Belongs to the dynein heavy chain family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L23195; AAA60323.1; -.
DR   EMBL; L25122; AAA28492.1; -.
DR   PIR; A54794; A54794.
DR   FlyBase; FBgn0010349; Dhc64C.
DR   GO; GO:0045169; C:fusome; NAS.
DR   GO; GO:0008567; F:dynein ATPase activity; IDA.
DR   GO; GO:0003777; F:microtubule motor activity; IDA.
DR   GO; GO:0007098; P:centrosome cycle; IMP.
DR   GO; GO:0045478; P:fusome organization and biogenesis; NAS.
DR   GO; GO:0007018; P:microtubule-based movement; IDA.
DR   GO; GO:0007294; P:oocyte determination; IMP.
DR   GO; GO:0007292; P:oogenesis; IMP.
DR   InterPro; IPR003593; AAA_ATPase.
DR   InterPro; IPR004273; Dynein_heavy.
DR   Pfam; PF03028; Dynein_heavy; 1.
DR   SMART; SM00382; AAA; 4.
KW   Motor protein; Dynein; Microtubules; ATP-binding; Repeat; Coiled coil.
FT   DOMAIN       1   1856       STEM (BY SIMILARITY).
FT   DOMAIN    1857   2084       AAA 1 (BY SIMILARITY).
FT   DOMAIN    2166   2437       AAA 2 (BY SIMILARITY).
FT   DOMAIN    2541   2790       AAA 3 (BY SIMILARITY).
FT   DOMAIN    2884   3153       AAA 4 (BY SIMILARITY).
FT   DOMAIN    3189   3478       STALK (BY SIMILARITY).
FT   DOMAIN    3539   3768       AAA 5 (BY SIMILARITY).
FT   DOMAIN    3989   4205       AAA 6 (BY SIMILARITY).
FT   DOMAIN     530    565       COILED COIL (POTENTIAL).
FT   DOMAIN     774    794       COILED COIL (POTENTIAL).
FT   DOMAIN    1264   1368       COILED COIL (POTENTIAL).
FT   DOMAIN    3189   3261       COILED COIL (POTENTIAL).
FT   DOMAIN    3382   3478       COILED COIL (POTENTIAL).
FT   DOMAIN    3723   3782       COILED COIL (POTENTIAL).
FT   NP_BIND   1895   1902       ATP (POTENTIAL).
FT   NP_BIND   2210   2217       ATP (POTENTIAL).
FT   NP_BIND   2580   2587       ATP (POTENTIAL).
FT   NP_BIND   2922   2929       ATP (POTENTIAL).
SQ   SEQUENCE   4639 AA;  530152 MW;  057A7D8800CCD07E CRC64;

  Query Match             70.8%;  Score 34;  DB 1;  Length 4639;
  Best Local Similarity   62.5%;  Pred. No. 3.8e+02;
  Matches    5;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 KWLIQLKK 9
              :|:|:|:|
Db       4088 QWLVQLEK 4095


RESULT 10
DYHC_DICDI
ID   DYHC_DICDI     STANDARD;      PRT;  4725 AA.
AC   P34036;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Dynein heavy chain, cytosolic (DYHC).
GN   DHCA.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AX3;
RX   MEDLINE=93107159; PubMed=1469051;
RA   Koonce M.P., Grissom P.M., McIntosh J.R.;
RT   "Dynein from Dictyostelium: primary structure comparisons between a
RT   cytoplasmic motor enzyme and flagellar dynein.";
RL   J. Cell Biol. 119:1597-1604(1992).
RN   [2]
RP   CHARACTERIZATION.
RX   MEDLINE=95170719; PubMed=7866389;
RA   Koonce M.P., Grissom P.M., Lyon M., Pope T., McIntosh J.R.;
RT   "Molecular characterization of a cytoplasmic dynein from
RT   Dictyostelium.";
RL   J. Eukaryot. Microbiol. 41:645-651(1994).
CC   -!- FUNCTION: Cytoplasmic dynein acts as a motor for the intracellular
CC       retrograde motility of vesicles and organelles along microtubules.
CC       Dynein has ATPase activity; the force-producing power stroke is
CC       thought to occur on release of ADP.
CC   -!- SUBUNIT: Consists of at least two heavy chains and a number of
CC       intermediate and light chains.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- DOMAIN: Dynein heavy chains probably consist of an N-terminal stem
CC       (which binds cargo and interacts with other dynein components),
CC       and the head or motor domain. The motor contains six tandemly-
CC       linked AAA domains in the head, which form a ring. A stalk-like
CC       structure (formed by two of the coiled coil domains) protrudes
CC       between AAA 4 and AAA 5 and terminates in a microtubule-binding
CC       site. A seventh domain may also contribute to this ring; it is not
CC       clear whether the N-terminus or the C-terminus forms this extra
CC       domain. There are four well-conserved and two non-conserved ATPase
CC       sites, one per AAA domain. Probably only one of these (within AAA
CC       1) actually hydrolyzes ATP, the others may serve a regulatory
CC       function.
CC   -!- SIMILARITY: Belongs to the dynein heavy chain family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z15124; CAA78827.1; -.
DR   PIR; A44357; A44357.
DR   DictyDb; DD01045; dhcA.
DR   InterPro; IPR003593; AAA_ATPase.
DR   InterPro; IPR004273; Dynein_heavy.
DR   Pfam; PF03028; Dynein_heavy; 1.
DR   SMART; SM00382; AAA; 3.
KW   Motor protein; Dynein; Microtubules; ATP-binding; Repeat; Coiled coil.
FT   DOMAIN       1   1930       STEM (BY SIMILARITY).
FT   DOMAIN    1931   2153       AAA 1 (BY SIMILARITY).
FT   DOMAIN    2233   2526       AAA 2 (BY SIMILARITY).
FT   DOMAIN    2630   2880       AAA 3 (BY SIMILARITY).
FT   DOMAIN    2973   3247       AAA 4 (BY SIMILARITY).
FT   DOMAIN    3266   3580       STALK (BY SIMILARITY).
FT   DOMAIN    3633   3862       AAA 5 (BY SIMILARITY).
FT   DOMAIN    4093   4307       AAA 6 (BY SIMILARITY).
FT   DOMAIN     864    897       COILED COIL (POTENTIAL).
FT   DOMAIN    1199   1219       COILED COIL (POTENTIAL).
FT   DOMAIN    1338   1367       COILED COIL (POTENTIAL).
FT   DOMAIN    1420   1436       COILED COIL (POTENTIAL).
FT   DOMAIN    1656   1684       COILED COIL (POTENTIAL).
FT   DOMAIN    2226   2248       COILED COIL (POTENTIAL).
FT   DOMAIN    2437   2457       COILED COIL (POTENTIAL).
FT   DOMAIN    3266   3344       COILED COIL (POTENTIAL).
FT   DOMAIN    3478   3580       COILED COIL (POTENTIAL).
FT   DOMAIN    3849   3876       COILED COIL (POTENTIAL).
FT   NP_BIND   1969   1976       ATP (POTENTIAL).
```

```
FT   NP_BIND    2271   2278       ATP (POTENTIAL).
FT   NP_BIND    2669   2676       ATP (POTENTIAL).
FT   NP_BIND    3011   3018       ATP (POTENTIAL).
SQ   SEQUENCE   4725 AA;  538807 MW;  9A0319915D579D3A CRC64;

  Query Match            70.8%;  Score 34;  DB 1;  Length 4725;
  Best Local Similarity  62.5%;  Pred. No. 3.8e+02;
  Matches    5;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 KWLIQLKK 9
              :||:|||:|
Db       4190 QWLVQLEK 4197


RESULT 11
NCCY_ALCXX
ID   NCCY_ALCXX     STANDARD;      PRT;    95 AA.
AC   Q44581;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   Nickel-cobalt-cadmium resistance protein nccY.
GN   NCCY.
OS   Alcaligenes xylosoxydans xylosoxydans (Achromobacter xylosoxidans).
OG   Plasmid pTOM9.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC   Alcaligenaceae; Achromobacter.
OX   NCBI_TaxID=515;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=31A;
RX   MEDLINE=95050278; PubMed=7961470;
RA   Schmidt T., Schlegel H.G.;
RT   "Combined nickel-cobalt-cadmium resistance encoded by the ncc locus
RT   of Alcaligenes xylosoxidans 31A.";
RL   J. Bacteriol. 176:7045-7054(1994).
CC   -!- FUNCTION: COMPONENT OF THE NCC CATION-EFFLUX SYSTEM THAT CONFERS
CC       RESISTANCE TO NICKEL, COBALT AND CADMIUM. MAY BE INVOLVED IN THE
CC       REGULATION OF NCC.
CC   -!- SIMILARITY: TO A.EUTROPHUS CNRY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L31363; AAA65101.1; -.
DR   PIR; I39575; I39575.
KW   Plasmid; Nickel; Cobalt; Cadmium.
SQ   SEQUENCE   95 AA;  10422 MW;  4524AA1F3EB37987 CRC64;

  Query Match            68.8%;  Score 33;  DB 1;  Length 95;
  Best Local Similarity  62.5%;  Pred. No. 14;
  Matches    5;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 KWLIQLKK 9
              :||:| ||
Db          6 EWLVQAKK 13


RESULT 12
RAC2_ARATH
ID   RAC2_ARATH     STANDARD;      PRT;    201 AA.
AC   Q38903;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   RAC-like GTP binding protein ARAC2.
GN   ARAC2 OR AT5G45970 OR MCL19.1.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=98009984; PubMed=9349271;
RA   Winge P., Brembu T., Bones A.M.;
RT   "Cloning and characterization of rac-like cDNAs from Arabidopsis
RT   thaliana.";
RL   Plant Mol. Biol. 35:483-495(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Landsberg erecta;
RA   Winge P., Brembu T., Kristensen R., Bones A.M.;
RL   Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=98069011; PubMed=9405937;
RA   Kotani H., Nakamura Y., Sato S., Kaneko T., Asamizu E., Miyajima N.,
RA   Tabata S.;
RT   "Structural analysis of Arabidopsis thaliana chromosome 5. II.
RT   Sequence features of the regions of 1,044,062 bp covered by thirteen
RT   physically assigned P1 clones.";
RL   DNA Res. 4:291-300(1997).
CC   -!- TISSUE SPECIFICITY: EXPRESSED EXCLUSIVELY IN THE ROOT, HYPOCOTYL
CC       AND STEM.
CC   -!- SIMILARITY: BELONGS TO THE SMALL GTPASE SUPERFAMILY. RHO FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U43026; AAC49852.1; -.
DR   EMBL; AF115469; AAF40241.1; -.
DR   EMBL; AB006698; BAB08242.1; -.
DR   PIR; T48862; T48862.
DR   HSSP; P21181; 1AM4.
DR   InterPro; IPR003578; GTPase_Rho.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00174; RHO; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
KW   GTP-binding; Prenylation; Lipoprotein.
FT   NP_BIND      13     20       GTP (BY SIMILARITY).
FT   NP_BIND      60     64       GTP (BY SIMILARITY).
FT   NP_BIND     118    121       GTP (BY SIMILARITY).
FT   DOMAIN       35     43       EFFECTOR REGION (POTENTIAL).
FT   LIPID       198    198       GERANYL-GERANYL (POTENTIAL).
SQ   SEQUENCE   201 AA;  22393 MW;  B3E73F55BDBB8028 CRC64;

  Query Match            68.8%;  Score 33;  DB 1;  Length 201;
  Best Local Similarity  75.0%;  Pred. No. 28;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLK 8
              |||| :||
Db         98 KKWLPELK 105


RESULT 13
V282_ARATH
ID   V282_ARATH     STANDARD;      PRT;    209 AA.
AC   O65421;
DT   16-OCT-2001 (Rel. 40, Created)
```

```
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   VPS28 protein homolog 2.
GN   AT4G21560 OR F18E5.180 OR F17L22.20.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=20083488; PubMed=10617198;
RA   Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
RA   Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,
RA   Harris B., Ansorge W., Brandt P., Grivell L.A., Rieger M.,
RA   Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M.,
RA   Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidtheini T.,
RA   Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
RA   Vos P., Hoheisel J., Zimmermann W., Wedler H., Ridley P.,
RA   Langham S.-A., McCullagh B., Bilham L., Robben J.,
RA   Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandenbussche F.,
RA   Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
RA   Weitzenegger T., Bothe G., Ramsperger U., Hilbert H., Braun M.,
RA   Holzer E., Brandt A., Peters S., van Staveren M., Dirkse W.,
RA   Mooijman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
RA   Berneiser S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
RA   De Keyser A., Buysshaert C., Gielen J., Villarroel R., De Clercq R.,
RA   Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
RA   Clark L., Doggett J., Hall S., Kay M., Lennard N., McLay K., Mayes R.,
RA   Pettett A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,
RA   Borkova D., Bloecker H., Scharfe M., Grimm M., Loehnert T.-H.,
RA   Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
RA   Gabel C., Fuchs M., Fartmann B., Granderath K., Dauner D., Herzl A.,
RA   Neumann S., Argiriou A., Vitale D., Liguori R., Piravandi E.,
RA   Massenet O., Quigley F., Clabauld G., Muendlein A., Felber R.,
RA   Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
RA   Chefdor F., Cooke R., Berger C., Monfort A., Casacuberta E.,
RA   Gibbons T., Weber N., Vandenbol M., Bargues M., Terol J., Torres A.,
RA   Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
RA   Heijnen L., Schwarz S., Scholler P., Heber S., Francs P., Bielke C.,
RA   Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
RA   Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
RA   Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
RA   Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
RA   Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
RA   Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
RA   Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
RA   Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,
RA   Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
RA   Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshu C.,
RA   Antonoiu B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
RA   Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,
RA   Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
RA   Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,
RA   Chen E., Marra M., Martienssen R., McCombie W.R.;
RT   "Sequence and analysis of chromosome 4 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 402:769-777(1999).
CC   -!- SIMILARITY: BELONGS TO THE VPS28 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AL022603; CAA18720.1; -.
DR   EMBL; AL035527; CAB36800.1; -.
DR   EMBL; AL161555; CAB81263.1; -.
DR   PIR; T05164; T05164.
DR   Pfam; PF03997; VPS28; 1.
KW   Hypothetical protein; Transport; Protein transport.
SQ   SEQUENCE   209 AA;  23494 MW;  8A8BB4B415D9786B CRC64;

  Query Match             68.8%;  Score 33;  DB 1;  Length 209;
  Best Local Similarity   55.6%;  Pred. No. 29;
  Matches    5;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |:|||::| |
Db        166 KEWLLRLSK 174


RESULT 14
GCH1_PASMU
ID   GCH1_PASMU     STANDARD;      PRT;   218 AA.
AC   P57865;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   GTP cyclohydrolase I (EC 3.5.4.16) (GTP-CH-I).
GN   FOLE OR PM0693.
OS   Pasteurella multocida.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Pasteurella.
OX   NCBI_TaxID=747;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Pm70;
RX   MEDLINE=21145866; PubMed=11248100;
RA   May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT   "Complete genomic sequence of Pasteurella multocida Pm70.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC   -!- CATALYTIC ACTIVITY: GTP + 2 H(2)O = formate + 2-amino-4-hydroxy-6-
CC       (erythro-1,2,3-trihydroxypropyl)dihydropteridine triphosphate.
CC   -!- PATHWAY: Tetrahydrofolate biosynthesis; first step.
CC   -!- SUBUNIT: Homopolymer (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE GTP CYCLOHYDROLASE I FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE006106; AAK02777.1; -.
DR   HSSP; P27511; 1A8R.
DR   HAMAP; MF_00223; -; 1.
DR   InterPro; IPR001474; GTP_cyclohydroI.
DR   Pfam; PF01227; GTP_cyclohydroI; 1.
DR   ProDom; PD003330; GTP_cyclohydroI; 1.
DR   TIGRFAMs; TIGR00063; folE; 1.
DR   PROSITE; PS00859; GTP_CYCLOHYDROL_1_1; 1.
DR   PROSITE; PS00860; GTP_CYCLOHYDROL_1_2; 1.
KW   One-carbon metabolism; Hydrolase; Complete proteome.
FT   DISULFID    109    180       BY SIMILARITY.
SQ   SEQUENCE   218 AA;  24953 MW;  12F1FC8129E790C9 CRC64;

  Query Match             68.8%;  Score 33;  DB 1;  Length 218;
  Best Local Similarity   66.7%;  Pred. No. 30;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||:| | |
Db        127 KKWVIGLSK 135


RESULT 15
EXOR_RHILV
ID   EXOR_RHILV     STANDARD;      PRT;   267 AA.
AC   Q52822;
```

```
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Exopolysaccharide production negative regulator precursor.
GN   EXOR.
OS   Rhizobium leguminosarum (biovar viciae).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.
OX   NCBI_TaxID=387;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WSM710 / WR6-35;
RX   MEDLINE=97346048; PubMed=9202471;
RA   Reeve W.G., Dilworth M.J., Tiwari R.P., Glenn A.R.;
RT   "Regulation of exopolysaccharide production in Rhizobium leguminosarum
RT   biovar viciae WSM710 involves exoR.";
RL   Microbiology 143:1951-1958(1997).
CC   -!- FUNCTION: NEGATIVELY MODULATES EXOPOLYSACCHARIDE (EPS)
CC       BIOSYNTHESIS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; L39937; AAB63941.1; -.
DR   InterPro; IPR006597; Sel_like.
DR   SMART; SM00671; SEL1; 3.
KW   Exopolysaccharide synthesis; Signal.
FT   SIGNAL        1     23       POTENTIAL.
FT   CHAIN        24    267       EXOPOLYSACCHARIDE PRODUCTION NEGATIVE
FT                                REGULATOR.
SQ   SEQUENCE   267 AA;  28922 MW;  E45E2E0538E1F573 CRC64;

  Query Match             68.8%;  Score 33;  DB 1;  Length 267;
  Best Local Similarity   66.7%;  Pred. No. 37;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||| | :|
Db        185 KKWLNQARK 193


Search completed: October 14, 2003, 14:34:08
Job time : 6.54167 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 14:33:18 ; Search time 29.8125 Seconds
(without alignments)
77.903 Million cell updates/sec

Title: US-09-881-490-164
Perfect score: 48
Sequence: 1 KKWLIQLKK 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	40	83.3	369	2	Q93CS6	Q93cs6 shigella bo
2	38	79.2	229	5	Q8MP31	Q8mp31 dictyosteli
3	37	77.1	189	16	Q8Z1K3	Q8z1k3 salmonella
4	37	77.1	275	16	Q9JYL8	Q9jyl8 neisseria m
5	36	75.0	89	2	Q9EYA7	Q9eya7 escherichia
6	36	75.0	278	4	Q9UJ05	Q9uj05 homo sapien
7	36	75.0	284	4	Q96BV3	Q96bv3 homo sapien
8	36	75.0	459	10	Q8H3G5	Q8h3q5 oryza sativ
9	36	75.0	578	10	O23488	O23488 arabidopsis
10	36	75.0	652	5	Q8I427	Q8i427 plasmodium
11	36	75.0	1566	5	Q9N3E9	Q9n3e9 caenorhabdi
12	35	72.9	136	16	Q8EI76	Q8ei76 shewanella
13	35	72.9	198	10	O65062	O65062 picea maria
14	35	72.9	261	16	Q9CNC3	Q9cnc3 pasteurella
15	35	72.9	294	8	Q8HIS8	Q8his8 monosiga br
16	35	72.9	338	5	O01679	O01679 bombyx mori
17	35	72.9	381	16	Q8DSD4	Q8dsd4 streptococc
18	35	72.9	585	16	O66759	O66759 aquifex aeo
19	35	72.9	643	5	Q8IIV6	Q8iiv6 plasmodium
20	35	72.9	731	10	Q9AW75	Q9aw75 guillardia
21	35	72.9	767	10	Q8GXD0	Q8gxd0 arabidopsis
22	35	72.9	2655	10	Q9C8A3	Q9c8a3 arabidopsis
23	34	70.8	59	16	Q8YW77	Q8yw77 anabaena sp
24	34	70.8	89	9	Q9XJI1	Q9xji1 bacteriopha
25	34	70.8	89	16	Q8XEH8	Q8xeh8 escherichia
26	34	70.8	115	16	Q8CV09	Q8cv09 oceanobacil
27	34	70.8	234	11	Q9D334	Q9d334 mus musculu
28	34	70.8	235	16	Q8XL88	Q8xl88 clostridium
29	34	70.8	245	4	O60573	O60573 homo sapien
30	34	70.8	245	11	O88503	O88503 mus musculu
31	34	70.8	245	11	Q8BMB3	Q8bmb3 mus musculu
32	34	70.8	276	16	Q9JTL7	Q9jtl7 neisseria m
33	34	70.8	285	16	O31436	O31436 bacillus su
34	34	70.8	287	16	Q98QY4	Q98qy4 mycoplasma
35	34	70.8	343	1	Q9HH38	Q9hh38 methanobact
36	34	70.8	432	2	Q8GCX1	Q8gcx1 enterococcu
37	34	70.8	521	5	Q9VPX7	Q9vpx7 drosophila
38	34	70.8	526	4	Q9Y3N2	Q9y3n2 homo sapien
39	34	70.8	547	16	Q8XE44	Q8xe44 escherichia
40	34	70.8	547	16	Q8FFQ4	Q8ffq4 escherichia
41	34	70.8	732	16	Q99XP7	Q99xp7 streptococc
42	34	70.8	732	16	Q8K5K3	Q8k5k3 streptococc
43	34	70.8	839	5	Q961M8	Q961m8 drosophila
44	34	70.8	954	4	Q8WXV8	Q8wxv8 homo sapien
45	34	70.8	957	4	Q9H0V3	Q9h0v3 homo sapien

ALIGNMENTS

RESULT 1
Q93CS6
ID Q93CS6 PRELIMINARY; PRT; 369 AA.
AC Q93CS6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Putative glycosyl transferase.
GN WBAT.
OS Shigella boydii.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=621;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21481970; PubMed=11598067;
RA Wang L., Qu W., Reeves P.R.;
RT "Sequence Analysis of Four Shigella boydii O-Antigen Loci: Implication
RT for Escherichia coli and Shigella Relationships.";
RL Infect. Immun. 69:6923-6930(2001).
DR EMBL; AF402314; AAL27338.1; -.
DR InterPro; IPR001296; Glyco_trans_1.
DR Pfam; PF00534; Glycos_transf_1; 1.
KW Transferase.
SQ SEQUENCE 369 AA; 42760 MW; 3E2CCE371A6BA0BB CRC64;

Query Match 83.3%; Score 40; DB 2; Length 369;
Best Local Similarity 77.8%; Pred. No. 23;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 KKWLIQLKK 9
              |||||:||:
Db        359 KKWLIKLKQ 367
```

RESULT 2
Q8MP31
ID Q8MP31 PRELIMINARY; PRT; 229 AA.
AC Q8MP31;

```
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AX4;
RA   Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
RA   Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA   Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT   "Sequence and Analysis of Chromosome 2 of Dictyostelium.";
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC117178; AAM44388.1; -.
DR   InterPro; IPR001064; Crystallin.
DR   PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   229 AA;  27735 MW;  FDC1E5C8B9647DD6 CRC64;

  Query Match             79.2%;  Score 38;  DB 5;  Length 229;
  Best Local Similarity   66.7%;  Pred. No. 34;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |||| ::|| |
Db        119 KKWLIEIKK 127


RESULT 3
Q8Z1K3
ID   Q8Z1K3      PRELIMINARY;      PRT;   189 AA.
AC   Q8Z1K3;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Putative membrane protein.
GN   STY4561.
OS   Salmonella typhi.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Salmonella.
OX   NCBI_TaxID=601;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CT18;
RX   MEDLINE=21534947; PubMed=11677608;
RA   Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA   Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
RA   Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
RA   Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA   Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA   Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA   Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA   Whitehead S., Barrell B.G.;
RT   "Complete genome sequence of a multiple drug resistant Salmonella
RT   enterica serovar Typhi CT18.";
RL   Nature 413:848-852(2001).
DR   EMBL; AL627282; CAD09337.1; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   189 AA;  22077 MW;  C817745ABC69BE2A CRC64;

  Query Match             77.1%;  Score 37;  DB 16;  Length 189;
  Best Local Similarity   66.7%;  Pred. No. 44;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              |:|||| |:
Db        101 KRWLIQAKR 109


RESULT 4
Q9JYL8
ID   Q9JYL8      PRELIMINARY;      PRT;   275 AA.
AC   Q9JYL8;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Phytoene synthase-related protein.
GN   NMB1521.
OS   Neisseria meningitidis (serogroup B).
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MC58 / Serogroup B;
RX   MEDLINE=20175755; PubMed=10710307;
RA   Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA   Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA   Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA   Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA   Mason T., Ciecko A., Parksey D.S., Blair E., Cittone H., Clark E.B.,
RA   Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
RA   Gill J., Scarlato V., Masignani V., Pizza M., Grandi G., Sun L.,
RA   Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT   "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT   MC58.";
RL   Science 287:1809-1815(2000).
DR   EMBL; AE002501; AAF41877.1; -.
DR   TIGR; NMB1521; -.
DR   InterPro; IPR002060; Squ/phyt_synthse.
DR   Pfam; PF00494; SQS_PSY; 1.
KW   Complete proteome.
SQ   SEQUENCE   275 AA;  30613 MW;  B8B2C2D77A549EA6 CRC64;

  Query Match             77.1%;  Score 37;  DB 16;  Length 275;
  Best Local Similarity   77.8%;  Pred. No. 62;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              | ||| |||
Db        263 KDWLIMLKK 271


RESULT 5
Q9EYA7
ID   Q9EYA7      PRELIMINARY;      PRT;    89 AA.
AC   Q9EYA7;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Kil protein.
GN   KIL.
OS   Escherichia coli O157:H7.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=83334;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O157:H7;
RX   MEDLINE=20564182; PubMed=11111050;
RA   Yokoyama K., Makino K., Kubota Y., Watanabe M., Kimura S.,
RA   Yutsudo C.H., Kurokawa K., Ishii K., Hattori M., Abe H., Iida T.,
RA   Yamamoto K., Hayashi T., Yasunaga T., Honda T., Sasakawa C.,
RA   Shinagawa H.;
RT   "Complete nucleotide sequence of the defective Sakai-VT1 prophage
RT   carrying the shiga toxin 1 genes of the enterohemorrhagic Escherichia
RT   coli O157:H7 derived from the Sakai outbreak.";
RL   Gene 258:127-139(2000).
DR   EMBL; AP000400; BAB19614.1; -.
SQ   SEQUENCE   89 AA;  10127 MW;  CD30904BF080D5FD CRC64;

  Query Match             75.0%;  Score 36;  DB 2;  Length 89;
  Best Local Similarity   55.6%;  Pred. No. 33;
  Matches    5;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy          1 KKWLIQLKK 9
              ||||: ||:
Db         74 KKWILMLKR 82


RESULT 6
Q9UJ05
ID   Q9UJ05      PRELIMINARY;      PRT;   278 AA.
AC   Q9UJ05;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   DJ486I3.1 (Novel protein) (Fragment).
GN   DJ486I3.1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Barlow K.;
RL   Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL050331; CAB55880.1; -.
FT   NON_TER       1       1
SQ   SEQUENCE   278 AA;  31677 MW;  0088AFE6B3BC3F6F CRC64;

  Query Match             75.0%;  Score 36;  DB 4;  Length 278;
  Best Local Similarity   87.5%;  Pred. No. 95;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLK 8
              |||| |||
Db         25 KKWLRQLK 32


RESULT 7
Q96BV3
ID   Q96BV3      PRELIMINARY;      PRT;   284 AA.
AC   Q96BV3;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Hypothetical protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Bladder;
RA   Strausberg R.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC015138; AAH15138.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   284 AA;  32370 MW;  BCD8F0A676995S0A CRC64;

  Query Match             75.0%;  Score 36;  DB 4;  Length 284;
  Best Local Similarity   87.5%;  Pred. No. 97;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLK 8
              |||| |||
Db         31 KKWLRQLK 38


RESULT 8
Q8H3G5
ID   Q8H3G5      PRELIMINARY;      PRT;   459 AA.
AC   Q8H3G5;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   OSJNBb0062D12.24 protein.
GN   OSJNBB0062D12.24.
OS   Oryza sativa (japonica cultivar-group).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=39947;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Nipponbare;
RA   Sasaki T., Matsumoto T., Katayose Y.;
RT   "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 7, BAC
RT   clone:OSJNBb0062D12.";
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AP005127; BAC16708.1; -.
SQ   SEQUENCE   459 AA;  51464 MW;  5D66AF12D91CF90F CRC64;

  Query Match             75.0%;  Score 36;  DB 10;  Length 459;
  Best Local Similarity   75.0%;  Pred. No. 1.5e+02;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 KWLIQLKK 9
              |||:| ||
Db        322 KWLVQKKK 329


RESULT 9
O23488
ID   O23488      PRELIMINARY;      PRT;   578 AA.
AC   O23488;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein (LET1 like protein).
GN   AT4G16440.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98121113; PubMed=9461215;
RA   Bevan M., Bancroft I., Bent E., Love K., Goodman H., Dean C.,
RA   Bergkamp R., Dirkse W., van Staveren M., Stiekema W., Drost L.,
RA   Ridley P., Hudson S.A., Patel K., Murphy G., Piffanelli P., Wedler H.,
RA   Wedler E., Wambutt R., Weitzenegger T., Pohl T.M., Terryn N.,
RA   Gielen J., Villarroel R., De Clerck R., van Montagu M., Lecharny A.,
RA   Auborg S., Gy I., Kreis M., Lao N., Kavanagh T., Hempel S., Kotter P.,
RA   Entian K.D., Rieger M., Schaeffer M., Funk B., Mueller-Auer S.,
RA   Silvey M., James R., Montfort A., Pons A., Puigdomenech P., Douka A.,
RA   Voukelatou E., Milioni D., Hatzopoulos P., Piravandi E., Obermaier B.,
RA   Hilbert H., Duesterhoft A., Moores T., Jones J.D.G., Eneva T.,
RA   Palme K., Benes V., Rechman S., Ansorge W., Cooke R., Berger C.,
RA   Delseny M., Voet M., Volckaert G., Mewes H.W., Klosterman S.,
RA   Schueller C., Chalwatzis N.;
RT   "Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of
RT   Arabidopsis thaliana.";
RL   Nature 391:485-488(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z97341; CAB10420.1; -.
DR   EMBL; AL161544; CAB78686.1; -.
DR   InterPro; IPR006459; DUF_A_thal_1569.
DR   InterPro; IPR003149; Fe_hyd_SSU.
DR   Pfam; PF02256; Fe_hyd_SSU; 1.
DR   TIGRFAMs; TIGR01569; A_tha_TIGR01569; 1.
SQ   SEQUENCE   578 AA;  64340 MW;  212385A45DC3C09A CRC64;

  Query Match             75.0%;  Score 36;  DB 10;  Length 578;
  Best Local Similarity   87.5%;  Pred. No. 1.9e+02;
```

```
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 KWLIQLKK 9
              ||| ||||
Db         30 KWLDQLKK 37


RESULT 10
Q8I427
ID   Q8I427      PRELIMINARY;      PRT;   652 AA.
AC   Q8I427;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Cell differentiation protein rcd1, putative.
GN   PFE0375W.
OS   Plasmodium falciparum (isolate 3D7).
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=36329;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Devlin K., Baker S., Davies P., Mungall K., Berriman M., Pain A.,
RA   Hall N., Bowman S., Churcher C., Quail M., Barrell B.;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22255708; PubMed=12368867;
RA   Hall N., Pain A., Berriman M., Churcher C., Harris B., Harris D.,
RA   Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,
RA   Buckee C.O., Burrows C., Cherevach I., Chillingworth C.,
RA   Chillingworth T., Christodoulou Z., Clark L., Clark R., Corton C.,
RA   Cronin A., Davies R., Davis P., Dear P., Dearden F., Doggett J.,
RA   Feltwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N., Hance Z.,
RA   Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P.,
RA   Humphray S., Jagels K., James K.D., Johnson D., Kerhornou A.,
RA   Knights A., Konfortov B., Kyes S., Larke N., Lawson D., Lennard N.,
RA   Line A., Maddison M., Mclean J., Mooney P., Moule S., Murphy L.,
RA   Oliver K., Ormond D., Price C., Quail M.A., Rabbinowitsch E.,
RA   Rajandream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M.,
RA   Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K.,
RA   Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,
RA   Sulston J.E., Craig A., Newbold C., Barrell B.G.;
RT   "Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13.";
RL   Nature 419:527-531(2002).
DR   EMBL; AL929351; CAD51441.1; -.
SQ   SEQUENCE   652 AA;  73423 MW;  90BBF62FBFB7FFFF CRC64;

  Query Match             75.0%;  Score 36;  DB 5;  Length 652;
  Best Local Similarity   85.7%;  Pred. No. 2.1e+02;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQL 7
              ||||:||
Db        482 KKWLLQL 488


RESULT 11
Q9N3E9
ID   Q9N3E9      PRELIMINARY;      PRT;  1566 AA.
AC   Q9N3E9;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical 180.6 kDa protein.
GN   Y54E10A.11.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RX   MEDLINE=99069613; PubMed=9851916;
RA   None;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology. The C. elegans Sequencing Consortium.";
RL   Science 282:2012-2018(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Wollam A., Becker M., Graves T., Hawkins M.;
RT   "The sequence of C. elegans cosmid Y54E10A.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Waterston R.;
RT   "Direct Submission.";
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC024810; AAF60772.2; -.
DR   WormPep; Y54E10A.11; CE29120.
DR   InterPro; IPR001841; Znf_ring.
DR   SMART; SM00184; RING; 1.
DR   PROSITE; PS50089; ZF_RING_2; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   1566 AA;  180607 MW;  F7D21D6EC9BA0F16 CRC64;

  Query Match             75.0%;  Score 36;  DB 5;  Length 1566;
  Best Local Similarity   85.7%;  Pred. No. 4.7e+02;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQL 7
              ||||:||
Db       1477 KKWLLQL 1483


RESULT 12
Q8EI76
ID   Q8EI76      PRELIMINARY;      PRT;   136 AA.
AC   Q8EI76;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein.
GN   SO0971.
OS   Shewanella oneidensis.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC   Alteromonadaceae; Shewanella.
OX   NCBI_TaxID=70863;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MR-1;
RX   MEDLINE=22297686; PubMed=12368813;
RA   Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA   Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,
RA   Meyer T., Tsapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,
RA   DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
RA   Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
RA   Vamathevan J., Weidman J., Impraim M., Lee K., Berry K., Lee C.,
RA   Mueller J., Khouri H., Gill J., Utterback T.R., McDonald L.A.,
RA   Feldblyum T.V., Smith H.O., Venter J.C., Nealson K.H., Fraser C.M.;
RT   "Genome sequence of the dissimilatory metal ion-reducing bacterium
RT   Shewanella oneidensis.";
RL   Nat. Biotechnol. 20:1118-1123(2002).
DR   EMBL; AE015542; AAN54045.1; -.
DR   TIGR; SO0971; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   136 AA;  15529 MW;  420710910979605C CRC64;

  Query Match             72.9%;  Score 35;  DB 16;  Length 136;
  Best Local Similarity   85.7%;  Pred. No. 75;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 KWLIQLK 8
              :||||||
Db         69 RWLIQLK 75
```

```
RESULT 13
O65062
ID   O65062      PRELIMINARY;      PRT:   198 AA.
AC   O65062;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Rac-like GTP binding protein.
GN   SB30.
OS   Picea mariana (Black spruce).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Picea.
OX   NCBI_TaxID=3335;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98278823; PubMed=9611216;
RA   Perry D.J., Bousquet J.;
RT   "Sequence-tagged-site (STS) markers of arbitrary genes. Development,
RT   characterization and analysis of linkage in black spruce.";
RL   Genetics 149:1089-1098(1998).
CC   -!- SIMILARITY: BELONGS TO THE SMALL GTPASE SUPERFAMILY. RHO FAMILY.
DR   EMBL; AF051223; AAC32124.1; -.
DR   HSSP; P21181; 1AM4.
DR   InterPro; IPR003578; GTPase_Rho.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00174; RHO; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
KW   GTP-binding; Lipoprotein; Prenylation.
SQ   SEQUENCE   198 AA;  21997 MW;  FA72F6EBB1D15E45 CRC64;

  Query Match             72.9%;  Score 35;  DB 10;  Length 198;
  Best Local Similarity   62.5%;  Pred. No. 1.1e-02;
  Matches    5;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLK 8
              ||||::|:
Db         98 KKWLLELR 105


RESULT 14
Q9CNC3
ID   Q9CNC3      PRELIMINARY;      PRT:   261 AA.
AC   Q9CNC3;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Hypothetical protein PM0509.
GN   PM0509.
OS   Pasteurella multocida.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Pasteurella.
OX   NCBI_TaxID=747;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Pm70;
RX   MEDLINE=21145866; PubMed=11248100;
RA   May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT   "Complete genomic sequence of Pasteurella multocida Pm70.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
DR   EMBL; AE006086; AAK02593.1; -.
DR   InterPro; IPR002654; Glyco_trans_25.
DR   Pfam; PF01755; Glyco_transf_25; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   261 AA;  30190 MW;  6D71C37C615EEF44 CRC64;

  Query Match             72.9%;  Score 35;  DB 16;  Length 261;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+02;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLKK 9
              :|||| |:|
Db          2 QKWLISLEK 10


RESULT 15
Q8HIS8
ID   Q8HIS8      PRELIMINARY;      PRT:   294 AA.
AC   Q8HIS8;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Orf294.
GN   ORF294.
OS   Monosiga brevicollis.
OG   Mitochondrion.
OC   Eukaryota; Choanoflagellida; Codonosigidae; Monosiga.
OX   NCBI_TaxID=81824;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20377911; PubMed=10916154;
RA   Bullerwell C.E., Burger G., Lang B.F.;
RT   "A novel motif for identifying rps3 homologs in fungal mitochondrial
RT   genomes.";
RL   Trends Biochem. Sci. 25:363-365(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Lang B.F., O'Kelly C.J., Nerad T.A., Gray M.W., Burger G.;
RT   "The closest unicellular relatives of animals.";
RL   Curr. Biol. 0:0-0(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Lang F.B., Bullerwell C.;
RL   Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Burger G., Lang B.F.;
RL   Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF538053; AAN28356.1; -.
KW   Mitochondrion.
SQ   SEQUENCE   294 AA;  34521 MW;  6CD1601F13C6D9D6 CRC64;

  Query Match             72.9%;  Score 35;  DB 8;  Length 294;
  Best Local Similarity   75.0%;  Pred. No. 1.5e+02;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKWLIQLK 8
              |||:| ||
Db         30 KKWIITLK 37


Search completed: October 14, 2003, 14:36:58
Job time : 29.8125 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 14:33:19 ; Search time 38.625 Seconds
(without alignments)
36.985 Million cell updates/sec

Title: US-09-881-490-191
Perfect score: 51
Sequence: 1 KCLIQLFCK 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	51	100.0	9	17	AAW04050	Antifungal peptide
2	51	100.0	9	18	AAW44590	Anti-fungal peptid
3	51	100.0	9	18	AAW44591	Anti-fungal peptid
4	51	100.0	9	20	AAY00567	Antifungal peptide
5	51	100.0	9	20	AAY00568	Antifungal peptide
6	51	100.0	9	22	AAB65491	Anti-fungal peptid
7	51	100.0	9	22	AAB65492	Anti-fungal peptid
8	51	100.0	11	17	AAW04051	Antifungal peptide
9	51	100.0	11	18	AAW44592	Anti-fungal peptid
10	51	100.0	11	20	AAY00569	Antifungal peptide
11	51	100.0	11	22	AAB65493	Anti-fungal peptid
12	39	76.5	316	22	AAG72962	Human olfactory re
13	39	76.5	731	22	AAM80128	Human protein SEQ
14	39	76.5	731	22	AAM41909	Human polypeptide
15	39	76.5	751	22	AAM40123	Human polypeptide
16	39	76.5	875	22	AAM79144	Human protein SEQ
17	39	76.5	886	24	AAE33484	Human REMAP-15 pro
18	39	76.5	944	24	ABP96125	Phoma betae aphidi
19	39	76.5	1600	22	AAE05958	Human phospholipas
20	39	76.5	1624	22	AAE05956	Human phospholipas
21	38	74.5	313	22	AAG72999	Olfactory receptor
22	37	72.5	37	23	ABG99477	Conus sp conotoxin
23	37	72.5	37	23	ABG99773	Conus sp conotoxin
24	37	72.5	56	22	AAU14900	Novel bone marrow
25	37	72.5	79	23	ABG99476	Conus sp conotoxin
26	37	72.5	217	22	AAG73054	Olfactory receptor
27	37	72.5	217	22	AAG73055	Olfactory receptor
28	37	72.5	217	22	AAG73057	Olfactory receptor
29	37	72.5	310	22	ABG03324	Novel human diagno
30	36	70.6	24	20	AAY45373	Human secreted pro
31	36	70.6	78	20	AAY11925	Human 5' EST secre
32	36	70.6	78	20	AAY13168	Human secreted pro
33	36	70.6	117	23	ABB89613	Human polypeptide
34	36	70.6	129	22	ABG07357	Novel human diagno
35	36	70.6	133	22	ABG22837	Novel human diagno
36	36	70.6	195	20	AAY45369	Human secreted pro
37	36	70.6	215	20	AAY40864	A human glutathion
38	36	70.6	215	21	AAB03805	Human glutathione-
39	36	70.6	215	22	AAB68866	Human glutathione-
40	36	70.6	243	23	AAE25017	Human drug metabol
41	36	70.6	262	22	AAM25798	Human protein sequ
42	36	70.6	476	22	ABB66584	Drosophila melanog
43	36	70.6	1855	20	AAY21804	Reading frame 1 pr
44	36	70.6	1855	21	AAY83272	Polypeptide encode
45	35	68.6	61	22	AAU45498	Propionibacterium

ALIGNMENTS

RESULT 1
AAW04050
ID AAW04050 standard; peptide; 9 AA.
XX
AC AAW04050;
XX
DT 04-NOV-1996 (first entry)
XX
DE Antifungal peptide XMP.360/XMP.361.
XX
KW Antifungal peptide; inhibitor; Domain III; polymorphonuclear leukocyte;
KW bactericidal/permeability-increasing protein; BPI; mammalian; PMN; fungi;
KW neutrophil; replication inhibitor; fungal infection; Aspergillus;
KW Cryptococcus; Candida; C.albicans; C.galabrat; C.krusei; C.lusitaniae;
KW C.parapsilosis; C.tropicalis; therapy.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 1
FT /note= "D-form residue"
FT Misc-difference 2
FT /note= "optionally D-form residue"
FT Modified-site 9
FT /note= "amidated D-form residue"
XX
PN WO9608509-A1.
XX
PD 21-MAR-1996.
XX
PF 20-JUL-1995; 95WO-US09262.
XX

```
PR   13-JAN-1995;   95US-0372105.
PR   15-SEP-1994;   94US-0306473.
XX
PA   (XOMA ) XOMA CORP.
XX
PI   Fadem MB,  Lim E,  Little RG;
XX
DR   WPI; 1996-179900/18.
XX
PT   Antifungal peptide(s) derived from Domain III of BPI protein  - used
PT   in vitro for killing or inhibiting replication of fungi, esp.
PT   Candida species
XX
PS   Claim 13; Page 167; 199pp; English.
XX
CC   AAW04000-W04160 represent antifungal peptides.  These sequences are
CC   based on (or derived from) Domain III of the
CC   bactericidal/permeability-increasing protein (BPI).  BPI is a protein
CC   that can be isolated from the granules of mammalian polymorphonuclear
CC   leukocytes (PMNs or neutrophils).  These antifungal peptides can be used
CC   for killing, or inhibiting replication of, fungi in vitro.  These
CC   sequences can also be used for treatment of a fungal infection involving
CC   fungi from the species Candida, Aspergillus and Cryptococcus.  The
CC   sequences are especially useful for treating C.albicans, C.galabrat,
CC   C.krusei, C.lusitaniae, C.parapsilosis and C.tropicalis infections.
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 51;  DB 17;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 9.3e-05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KCLIQLFCK 9
              |||||||||
Db          1 KCLIQLFCK 9


RESULT 2
AAW44590
ID   AAW44590 standard; peptide; 9 AA.
XX
AC   AAW44590;
XX
DT   27-APR-1998  (first entry)
XX
DE   Anti-fungal peptide #191 based on BPI protein (residues 142-169).
XX
KW   Anti-fungal peptide; bactericidal-permeability-increasing protein; BPI;
KW   polymorphonuclear leukocyte; fungicide.
XX
OS   Synthetic.
OS   Mammalia.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1
FT                   /note= "D-form residue"
FT   Misc-difference 2
FT                   /note= "D-form residue"
FT   Modified-site   9
FT                   /note= "C-terminal amide, D-form residue"
XX
PN   WO9704008-A1.
XX
PD   06-FEB-1997.
XX
PF   21-MAR-1996;   96WO-US03845.
XX
PR   20-JUL-1995;   95US-0504841.
XX
PA   (XOMA ) XOMA CORP.
XX
PI   Fadem MB,  Lim E,  Little RG;
XX
DR   WPI; 1997-132578/12.
XX
PT   Anti-fungal peptide(s) derived from or based on domain III of
PT   bactericidal-permeability-increasing protein - are used in vitro or
PT   in vivo as a fungicides
XX
PS   Claim 1; Page 204; 230pp; English.
XX
CC   This is a specifically claimed anti-fungal peptide which is based on
CC   domain III (amino acids 142-160) of bactericidal-permeability-increasing
CC   protein (BPI), isolated from the granules of mammalian polymorphonuclear
CC   leukocytes. It is used in compositions with diluents, carriers or
CC   adjuvants to treat fungal infections in patients. It may also be used in
CC   vitro to kill or inhibit the replication of fungi, such as in
CC   decontaminating fluids and sterilising medical and implant devices.
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 51;  DB 18;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 9.3e-05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KCLIQLFCK 9
              |||||||||
Db          1 KCLIQLFCK 9


RESULT 3
AAW44591
ID   AAW44591 standard; peptide; 9 AA.
XX
AC   AAW44591;
XX
DT   27-APR-1998  (first entry)
XX
DE   Anti-fungal peptide #192 based on BPI protein (residues 142-169).
XX
KW   Anti-fungal peptide; bactericidal-permeability-increasing protein; BPI;
KW   polymorphonuclear leukocyte; fungicide.
XX
OS   Synthetic.
OS   Mammalia.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1
FT                   /note= "D-form residue"
FT   Modified-site   9
FT                   /note= "C-terminal amide, D-form residue"
XX
PN   WO9704008-A1.
XX
PD   06-FEB-1997.
XX
PF   21-MAR-1996;   96WO-US03845.
XX
PR   20-JUL-1995;   95US-0504841.
XX
PA   (XOMA ) XOMA CORP.
XX
PI   Fadem MB,  Lim E,  Little RG;
XX
DR   WPI; 1997-132578/12.
XX
PT   Anti-fungal peptide(s) derived from or based on domain III of
PT   bactericidal-permeability-increasing protein - are used in vitro or
PT   in vivo as a fungicides
XX
PS   Claim 1; Page 204; 230pp; English.
XX
CC   This is a specifically claimed anti-fungal peptide which is based on
CC   domain III (amino acids 142-160) of bactericidal-permeability-increasing
CC   protein (BPI), isolated from the granules of mammalian polymorphonuclear
CC   leukocytes. It is used in compositions with diluents, carriers or
```

CC adjuvants to treat fungal infections in patients. It may also be used in
CC vitro to kill or inhibit the replication of fungi, such as in
CC decontaminating fluids and sterilising medical and implant devices.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 51; DB 18; Length 9;
Best Local Similarity 100.0%; Pred. No. 9.3e-05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KCLIQLFCK 9
||||||||| |
Db 1 KCLIQLFCK 9

RESULT 4
AAY00567
ID AAY00567 standard; Peptide; 9 AA.
XX
AC AAY00567;
XX
DT 07-MAY-1999 (first entry)
XX
DE Antifungal peptide XMP.360.
XX
KW Antifungal; BPI; bactericidal/permeability increasing protein;
KW Candida infection.
XX
OS Synthetic.
XX
PN US5858974-A.
XX
PD 12-JAN-1999.
XX
PF 21-MAR-1996; 96US-0621259.
XX
PR 21-MAR-1996; 96US-0621259.
PR 20-JUL-1995; 95US-0504841.
XX
PA (XOMA) XOMA CORP.
XX
PI Fadem MB, Lim E, Little RG;
XX
DR WPI; 1999-119956/10.
XX
PT Antifungal peptides - comprising part of bactericidal or
PT permeability-increasing protein sequence or related sequence
XX
PS Disclosure; Columns 185-186; 132pp; English.
XX
CC New peptides are provided which are based on Domain III (amino acids
CC 142-169) of human bactericidal/permeability-increasing protein (BPI).
CC The peptides all have a C-terminal amide. More particularly, the Claims
CC relate to: (1) a peptide that has an amino acid sequence of human BPI
CC from position 148 to position 161 (KSKVGWLIQLFHKK) and variants of the
CC sequence having antifungal activity; and (2) an antifungal peptide
CC having 6-14 amino acids comprising (a) a core sequence selected from
CC LIQL, IQLF, WLIQL, LIQLF and WLIQLF and (b) one or more cationic amino
CC acids selected from K, R, H, Orn (ornithine) and Dab (diaminobutyric
CC acid) at the N and/or C terminus of the core sequence. The new peptides
CC are used for killing or inhibiting replication of fungi in vitro; and
CC for treating fungal infections in vivo, in particular infections of
CC Candida, Aspergillus or Cryptococcus spp., especially C. albicans, C.
CC krusei, C. lusitaniae, C. parapsilosis or C. tropicalis. The peptide
CC can be administered topically, intravenously, orally or as an aerosol,
CC optionally together with a non-peptide antifungal agent.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 51; DB 20; Length 9;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KCLIQLFCK 9
||||||||| |
Db 1 KCLIQLFCK 9

RESULT 5
AAY00568
ID AAY00568 standard; Peptide; 9 AA.
XX
AC AAY00568;
XX
DT 07-MAY-1999 (first entry)
XX
DE Antifungal peptide XMP.361.
XX
KW Antifungal; BPI; bactericidal/permeability increasing protein;
KW Candida infection.
XX
CS Synthetic.
XX
PN US5858974-A.
XX
PD 12-JAN-1999.
XX
PF 21-MAR-1996; 96US-0621259.
XX
PR 21-MAR-1996; 96US-0621259.
PR 20-JUL-1995; 95US-0504841.
XX
PA (XOMA) XOMA CORP.
XX
PI Fadem MB, Lim E, Little RG;
XX
DR WPI; 1999-119956/10.
XX
PT Antifungal peptides - comprising part of bactericidal or
PT permeability-increasing protein sequence or related sequence
XX
PS Disclosure; Columns 187-188; 132pp; English.
XX
CC New peptides are provided which are based on Domain III (amino acids
CC 142-169) of human bactericidal/permeability-increasing protein (BPI).
CC The peptides all have a C-terminal amide. More particularly, the Claims
CC relate to: (1) a peptide that has an amino acid sequence of human BPI
CC from position 148 to position 161 (KSKVGWLIQLFHKK) and variants of the
CC sequence having antifungal activity; and (2) an antifungal peptide
CC having 6-14 amino acids comprising (a) a core sequence selected from
CC LIQL, IQLF, WLIQL, LIQLF and WLIQLF and (b) one or more cationic amino
CC acids selected from K, R, H, Orn (ornithine) and Dab (diaminobutyric
CC acid) at the N and/or C terminus of the core sequence. The new peptides
CC are used for killing or inhibiting replication of fungi in vitro; and
CC for treating fungal infections in vivo, in particular infections of
CC Candida, Aspergillus or Cryptococcus spp., especially C. albicans, C.
CC krusei, C. lusitaniae, C. parapsilosis or C. tropicalis. The peptide
CC can be administered topically, intravenously, orally or as an aerosol,
CC optionally together with a non-peptide antifungal agent.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 51; DB 20; Length 9;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KCLIQLFCK 9
||||||||| |
Db 1 KCLIQLFCK 9

RESULT 6
AAB65491
ID AAB65491 standard; Peptide; 9 AA.
XX
AC AAB65491;

XX
DT 27-MAR-2001 (first entry)
XX
DE Anti-fungal peptide XMP.360.
XX
KW Human; BPI; antifungal; polymorphonuclear leukocyte; neutrophil;
KW bactericidal/permeability-increasing protein; bactericidal;
KW fungal infection.
XX
OS Homo sapiens.
XX
PN US6156730-A.
XX
PD 05-DEC-2000.
XX
PF 08-JAN-1999; 99US-0227659.
XX
PR 21-MAR-1996; 96US-0621259.
PR 12-MAR-1993; 93US-0030644.
PR 15-JUL-1993; 93US-0093202.
PR 14-JAN-1994; 94US-0183222.
PR 11-MAR-1994; 94US-0209762.
PR 11-JUL-1994; 94US-0273540.
PR 15-SEP-1994; 94US-0306473.
PR 13-JAN-1995; 95US-0372105.
PR 20-JUL-1995; 95US-0504841.
XX
PA (XOMA) XOMA CORP.
XX
PI Lim E, Fadem MB, Little RG;
XX
DR WPI; 2001-090160/10.
XX
PT Novel anti-fungal peptides derived from domain III of
PT bactericidal/permeability-increasing protein useful for killing or
PT inhibiting replication of fungi and for treating fungal infections -
XX
PS Example 2; Columns 191-192; 134pp; English.
XX
CC The present invention relates to antifungal peptides (see
CC AAB65301-B65550) derived from Domain III (amino acids 142-169) of
CC bactericidal/permeability-increasing protein (BPI). The present sequence
CC is one such antifungal peptide. BPI is a protein isolated from the
CC granules of mammalian polymorphonuclear leukocytes (PMNs or
CC neutrophils). BPI has potent bactericidal activity against a broad range
CC of gram-negative bacteria. The peptides of the present invention are
CC useful for killing or inhibiting replication of fungi, and treating
CC infections caused by fungus belonging to Candida, Aspergillus,
CC Cryptococcus species such as C.albicans, C.glabrata, C.krusei,
CC C.lusitaniae, C.parapsilosis and C.tropicalis.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 51; DB 22; Length 9;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KCLIQLFCK 9
|||||||||
Db 1 KCLIQLFCK 9

RESULT 7
AAB65492
ID AAB65492 standard; Peptide; 9 AA.
XX
AC AAB65492;
XX
DT 27-MAR-2001 (first entry)
XX
DE Anti-fungal peptide XMP.361.
XX
KW Human; BPI; antifungal; polymorphonuclear leukocyte; neutrophil;
KW bactericidal/permeability-increasing protein; bactericidal;
KW fungal infection.
XX
OS Homo sapiens.
XX
PN US6156730-A.
XX
PD 05-DEC-2000.
XX
PF 08-JAN-1999; 99US-0227659.
XX
PR 21-MAR-1996; 96US-0621259.
PR 12-MAR-1993; 93US-0030644.
PR 15-JUL-1993; 93US-0093202.
PR 14-JAN-1994; 94US-0183222.
PR 11-MAR-1994; 94US-0209762.
PR 11-JUL-1994; 94US-0273540.
PR 15-SEP-1994; 94US-0306473.
PR 13-JAN-1995; 95US-0372105.
PR 20-JUL-1995; 95US-0504841.
XX
PA (XOMA) XOMA CORP.
XX
PI Lim E, Fadem MB, Little RG;
XX
DR WPI; 2001-090160/10.
XX
PT Novel anti-fungal peptides derived from domain III of
PT bactericidal/permeability-increasing protein useful for killing or
PT inhibiting replication of fungi and for treating fungal infections -
XX
PS Example 2; Columns 193-194; 134pp; English.
XX
CC The present invention relates to antifungal peptides (see
CC AAB65301-B65550) derived from Domain III (amino acids 142-169) of
CC bactericidal/permeability-increasing protein (BPI). The present sequence
CC is one such antifungal peptide. BPI is a protein isolated from the
CC granules of mammalian polymorphonuclear leukocytes (PMNs or
CC neutrophils). BPI has potent bactericidal activity against a broad range
CC of gram-negative bacteria. The peptides of the present invention are
CC useful for killing or inhibiting replication of fungi, and treating
CC infections caused by fungus belonging to Candida, Aspergillus,
CC Cryptococcus species such as C.albicans, C.glabrata, C.krusei,
CC C.lusitaniae, C.parapsilosis and C.tropicalis.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 51; DB 22; Length 9;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KCLIQLFCK 9
|||||||||
Db 1 KCLIQLFCK 9

RESULT 8
AAW04051
ID AAW04051 standard; peptide; 11 AA.
XX
AC AAW04051;
XX
DT 04-NOV-1996 (first entry)
XX
DE Antifungal peptide XMP.362.
XX
KW Antifungal peptide; inhibitor; Domain III; polymorphonuclear leukocyte;
KW bactericidal/permeability-increasing protein; BPI; mammalian; PMN; fungi;
KW neutrophil; replication inhibitor; fungal infection; Aspergillus;
KW Cryptococcus; Candida; C.albicans; C.galabrat; C.krusei; C.lusitaniae;
KW C.parapsilosis; C.tropicalis; therapy.
XX
OS Synthetic.

```
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1
FT                   /note= "D-form residue"
FT   Modified-site   11
FT                   /note= "amidated D-form residue"
XX
PN   WO9608509-A1.
XX
PD   21-MAR-1996.
XX
PF   20-JUL-1995;    95WO-US09262.
XX
PR   13-JAN-1995;    95US-0372105.
PR   15-SEP-1994;    94US-0306473.
XX
PA   (XOMA ) XOMA CORP.
XX
PI   Fadem MB,  Lim E,  Little RG;
XX
DR   WPI; 1996-179900/18.
XX
PT   Antifungal peptide(s) derived from Domain III of BPI protein  - used
PT   in vitro for killing or inhibiting replication of fungi, esp.
PT   Candida species
XX
PS   Claim 13; Page 168; 199pp; English.
XX
CC   AAW04000-W04160 represent antifungal peptides.  These sequences are
CC   based on (or derived from) Domain III of the
CC   bactericidal/permeability-increasing protein (BPI).  BPI is a protein
CC   that can be isolated from the granules of mammalian polymorphonuclear
CC   leukocytes (PMNs or neutrophils).  These antifungal peptides can be used
CC   for killing, or inhibiting replication of, fungi in vitro.  These
CC   sequences can also be used for treatment of a fungal infection involving
CC   fungi from the species Candida, Aspergillus and Cryptococcus.  The
CC   sequences are especially useful for treating C.albicans, C.galabrat,
CC   C.krusei, C.lusitaniae, C.parapsilosis and C.tropicalis infections.
XX
SQ   Sequence   11 AA;

  Query Match             100.0%;  Score 51;  DB 17;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 0.033;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KCLIQLFCK 9
              |||||||||
Db          2 KCLIQLFCK 10


RESULT 9
AAW44592
ID   AAW44592 standard; peptide; 11 AA.
XX
AC   AAW44592;
XX
DT   27-APR-1998  (first entry)
XX
DE   Anti-fungal peptide #193 based on BPI protein (residues 142-169).
XX
KW   Anti-fungal peptide; bactericidal-permeability-increasing protein; BPI;
KW   polymorphonuclear leukocyte; fungicide.
XX
OS   Synthetic.
OS   Mammalia.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1
FT                   /note= "D-form residue"
FT   Modified-site   11
FT                   /note= "C-terminal amide, D-form residue"
XX
PN   WO9704008-A1.
XX
PD   06-FEB-1997.
XX
PF   21-MAR-1996;    96WO-US03845.
XX
PR   20-JUL-1995;    95US-0504841.
XX
PA   (XOMA ) XOMA CORP.
XX
PI   Fadem MB,  Lim E,  Little RG;
XX
DR   WPI; 1997-132578/12.
XX
PT   Anti-fungal peptide(s) derived from or based on domain III of
PT   bactericidal-permeability-increasing protein - are used in vitro or
PT   in vivo as a fungicides
XX
PS   Claim 1; Page 205; 230pp; English.
XX
CC   This is a specifically claimed anti-fungal peptide which is based on
CC   domain III (amino acids 142-160) of bactericidal-permeability-increasing
CC   protein (BPI), isolated from the granules of mammalian polymorphonuclear
CC   leukocytes. It is used in compositions with diluents, carriers or
CC   adjuvants to treat fungal infections in patients. It may also be used in
CC   vitro to kill or inhibit the replication of fungi, such as in
CC   decontaminating fluids and sterilising medical and implant devices.
XX
SQ   Sequence   11 AA;

  Query Match             100.0%;  Score 51;  DB 18;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 0.033;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KCLIQLFCK 9
              |||||||||
Db          2 KCLIQLFCK 10


RESULT 10
AAY00569
ID   AAY00569 standard; Peptide; 11 AA.
XX
AC   AAY00569;
XX
DT   07-MAY-1999  (first entry)
XX
DE   Antifungal peptide XMP.362.
XX
KW   Antifungal; BPI; bactericidal/permeability increasing protein;
KW   Candida infection.
XX
OS   Synthetic.
XX
PN   US5858974-A.
XX
PD   12-JAN-1999.
XX
PF   21-MAR-1996;    96US-0621259.
XX
PR   21-MAR-1996;    96US-0621259.
PR   20-JUL-1995;    95US-0504841.
XX
PA   (XOMA ) XOMA CORP.
XX
PI   Fadem MB,  Lim E,  Little RG;
XX
DR   WPI; 1999-119956/10.
XX
PT   Antifungal peptides - comprising part of bactericidal or
PT   permeability-increasing protein sequence or related sequence
XX
PS   Disclosure; Columns 187-188; 132pp; English.
XX
```

```
CC   New peptides are provided which are based on Domain III (amino acids
CC   142-169) of human bactericidal/permeability-increasing protein (BPI).
CC   The peptides all have a C-terminal amide. More particularly, the Claims
CC   relate to: (1) a peptide that has an amino acid sequence of human BPI
CC   from position 148 to position 161 (KSKVGWLVQLFHKK) and variants of the
CC   sequence having antifungal activity; and (2) an antifungal peptide
CC   having 6-14 amino acids comprising (a) a core sequence selected from
CC   LIQL, IQLF, WLIQL, LIQLF and WLIQLF and (b) one or more cationic amino
CC   acids selected from K, R, H, Orn (ornithine) and Dab (diaminobutyric
CC   acid) at the N and/or C terminus of the core sequence. The new peptides
CC   are used for killing or inhibiting replication of fungi in vitro; and
CC   for treating fungal infections in vivo, in particular infections of
CC   Candida, Aspergillus or Cryptococcus spp., especially C. albicans, C.
CC   krusei, C. lusitaniae, C. parapsilosis or C. tropicalis. The peptide
CC   can be administered topically, intravenously, orally or as an aerosol,
CC   optionally together with a non-peptide antifungal agent.
XX
SQ   Sequence   11 AA;

  Query Match             100.0%;  Score 51;  DB 20;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 0.033;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KCLIQLFCK 9
              |||||||||
Db          2 KCLIQLFCK 10


RESULT 11
AAB65493
ID   AAB65493 standard; Peptide; 11 AA.
XX
AC   AAB65493;
XX
DT   27-MAR-2001  (first entry)
XX
DE   Anti-fungal peptide XMP.362.
XX
KW   Human; BPI; antifungal; polymorphonuclear leukocyte; neutrophil;
KW   bactericidal/permeability-increasing protein; bactericidal;
KW   fungal infection.
XX
OS   Homo sapiens.
XX
PN   US6156730-A.
XX
PD   05-DEC-2000.
XX
PF   08-JAN-1999;   99US-0227659.
XX
PR   21-MAR-1996;   96US-0621259.
PR   12-MAR-1993;   93US-0030644.
PR   15-JUL-1993;   93US-0093202.
PR   14-JAN-1994;   94US-0183222.
PR   11-MAR-1994;   94US-0209762.
PR   11-JUL-1994;   94US-0273540.
PR   15-SEP-1994;   94US-0306473.
PR   13-JAN-1995;   95US-0372105.
PR   20-JUL-1995;   95US-0504841.
XX
PA   (XOMA ) XOMA CORP.
XX
PI   Lim E,  Fadem MB,  Little RG;
XX
DR   WPI; 2001-090160/10.
XX
PT   Novel anti-fungal peptides derived from domain III of
PT   bactericidal/permeability-increasing protein useful for killing or
PT   inhibiting replication of fungi and for treating fungal infections -
XX
PS   Example 2; Columns 193-194; 134pp; English.
XX
CC   The present invention relates to antifungal peptides (see
CC   AAB65301-B65550) derived from Domain III (amino acids 142-169) of
CC   bactericidal/permeability-increasing protein (BPI). The present sequence
CC   is one such antifungal peptide. BPI is a protein isolated from the
CC   granules of mammalian polymorphonuclear leukocytes (PMNs or
CC   neutrophils). BPI has potent bactericidal activity against a broad range
CC   of gram-negative bacteria. The peptides of the present invention are
CC   useful for killing or inhibiting replication of fungi, and treating
CC   infections caused by fungus belonging to Candida, Aspergillus,
CC   Cryptococcus species such as C.albicans, C.glabrata, C.krusei,
CC   C.lusitaniae, C.parapsilosis and C.tropicalis.
XX
SQ   Sequence   11 AA;

  Query Match             100.0%;  Score 51;  DB 22;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 0.033;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KCLIQLFCK 9
              |||||||||
Db          2 KCLIQLFCK 10


RESULT 12
AAG72962
ID   AAG72962 standard; Protein; 316 AA.
XX
AC   AAG72962;
XX
DT   30-JUL-2001  (first entry)
XX
DE   Human olfactory receptor data exploratorium sequence, SEQ ID NO: 2644.
XX
KW   Human; olfactory receptor; OR; primary scent determination;
KW   secondary scent determination; polypeptide library; odour receptor;
KW   scent profile; scent fingerprint; scent representation;
KW   human olfactory receptor data exploratorium; HORDE.
XX
OS   Homo sapiens.
XX
PN   WO200127158-A2.
XX
PD   19-APR-2001.
XX
PF   06-OCT-2000; 2000WO-US27582.
XX
PR   08-OCT-1999;   99US-0158615.
PR   24-FEB-2000; 2000US-0184809.
XX
PA   (DIGI-) DIGISCENTS.
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Bellenson J,  Smith D,  Lancet D,  Glusman G,  Fuchs T,  Yanai I;
XX
DR   WPI; 2001-290713/30.
XX
PT   New polynucleotides which encode polypeptides involved in olfactory
PT   sensation for identifying olfactory agonists and antagonists -
XX
PS   Example 6; Page 1787-1788; 1857pp; English.
XX
CC   The present sequence is a polypeptide from the human olfactory receptor
CC   data exploratorium (HORDE). It was used as a query sequence in a database
CC   search of olfactory receptor (OR)-like sequences. The invention relates
CC   to isolated polynucleotides encoding polypeptides involved in olfactory
CC   sensation. The polynucleotides can be used in screening for olfactory
CC   agonists and antagonists. The methods allow for the determination of
CC   primary scents and the identification of the odour receptors used to
CC   detect these primary scents. The methods also enable determination of
CC   secondary scents and the identification of combinations of odour
CC   receptors that are involved in detecting such secondary scents. This
CC   enables the construction of a scent representation (also called a scent
CC   fingerprint or scent profile), which may be used to re-create and edit
CC   scents. Libraries of olfactory receptors are useful for determining the
```

```
CC   interaction pattern of a composition with the receptors, and can be
CC   used for determining differences in the olfactory faculties of different
CC   individuals.
XX
SQ   Sequence   316 AA;

  Query Match             76.5%;  Score 39;  DB 22;  Length 316;
  Best Local Similarity   85.7%;  Pred. No. 70;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 CLIQLFC 8
              |||||:|
Db         99 CLIQMFC 105


RESULT 13
AAM80128
ID   AAM80128 standard; Protein; 731 AA.
XX
AC   AAM80128;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human protein SEQ ID NO 3774.
XX
KW   Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW   vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW   tissue growth factor; immunomodulatory; cancer; leukaemia;
KW   nervous system disorder; arthritis; inflammation.
XX
OS   Homo sapiens.
XX
PN   WO200157190-A2.
XX
PD   09-AUG-2001.
XX
PF   05-FEB-2001; 2001WO-US04098.
XX
PR   03-FEB-2000; 2000US-0496914.
PR   27-APR-2000; 2000US-0560875.
PR   20-JUN-2000; 2000US-0598075.
PR   19-JUL-2000; 2000US-0620325.
PR   01-SEP-2000; 2000US-0654936.
PR   15-SEP-2000; 2000US-0663561.
PR   20-OCT-2000; 2000US-0693325.
PR   30-NOV-2000; 2000US-0728422.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Drmanac RT,  Asundi V,  Zhou P,  Xu C,  Cao Y,  Ma Y;
PI   Zhao QA,  Wang D,  Wang J,  Zhang J,  Ren F,  Chen R,  Wang ZW;
PI   Xue AJ,  Yang Y,  Wehrman T,  Goodrich R;
XX
DR   WPI; 2001-476283/51.
DR   N-PSDB; AAK53261.
XX
PT   Nucleic acids encoding polypeptides with cytokine-like activities,
PT   useful in diagnosis and gene therapy -
XX
PS   Claim 20; Page 432-433; 6221pp; English.
XX
CC   The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC   encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
CC   cytokine, cell proliferation or cell differentiation or which may induce
CC   production of other cytokines in other cell populations. The
CC   polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC   peptide therapy. The polypeptides have various cytokine-like activities,
CC   e.g. stem cell growth factor activity, haematopoiesis regulating
CC   activity, tissue growth factor activity, immunomodulatory activity and
CC   activin/inhibin activity and may be useful in the diagnosis and/or
CC   treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC   inflammation.
CC   Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC   (AAM80020) are omitted as the relevant pages from the sequence listing
CC   were missing at the time of publication.
XX
SQ   Sequence   731 AA;

  Query Match             76.5%;  Score 39;  DB 22;  Length 731;
  Best Local Similarity   75.0%;  Pred. No. 1.5e+02;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 CLIQLFCK 9
              |||||:|:
Db        466 CLIQLYCE 473


RESULT 14
AAM41909
ID   AAM41909 standard; Protein; 731 AA.
XX
AC   AAM41909;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polypeptide SEQ ID NO 6840.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US34263.
XX
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
PR   09-JUL-2000; 2000US-0598042.
PR   19-JUL-2000; 2000US-0620312.
PR   03-AUG-2000; 2000US-0653450.
PR   14-SEP-2000; 2000US-0662191.
PR   19-OCT-2000; 2000US-0693036.
PR   29-NOV-2000; 2000US-0727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Asundi V,  Chen R,  Ma Y,  Qian XB,  Ren F,  Wang D;
PI   Wang J,  Wang Z,  Wehrman T,  Xu C,  Xue AJ,  Yang Y,  Zhang J;
PI   Zhao QA,  Zhou P,  Goodrich R,  Drmanac RT;
XX
DR   WPI; 2001-442253/47.
DR   N-PSDB; AAI61065.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders
PT   such as central nervous system injuries -
XX
PS   Example 2; SEQ ID NO 6840; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC   the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
```

```
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification.
XX
SQ   Sequence   731 AA;

  Query Match             76.5%;  Score 39;  DB 22;  Length 731;
  Best Local Similarity   75.0%;  Pred. No. 1.5e+02;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 CLIQLFCK 9
              ||||||:|:
Db        466 CLIQLYCE 473


RESULT 15
AAM40123
ID   AAM40123 standard; Protein; 751 AA.
XX
AC   AAM40123;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polypeptide SEQ ID NO 3268.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US34263.
XX
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
PR   09-JUL-2000; 2000US-0598042.
PR   19-JUL-2000; 2000US-0620312.
PR   03-AUG-2000; 2000US-0633450.
PR   14-SEP-2000; 2000US-0662191.
PR   19-OCT-2000; 2000US-0693036.
PR   29-NOV-2000; 2000US-0727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Asundi V,  Chen R,  Ma Y,  Qian XB,  Ren F,  Wang D;
PI   Wang J,  Wang Z,  Wehrman T,  Xu C,  Xue AJ,  Yang Y,  Zhang J;
PI   Zhao QA,  Zhou P,  Goodrich R,  Drmanac RT;
XX
DR   WPI; 2001-442253/47.
DR   N-PSDB; AAI59279.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders
PT   such as central nervous system injuries -
XX
PS   Example 5; SEQ ID NO 3268; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC   the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification.
XX
SQ   Sequence   751 AA;

  Query Match             76.5%;  Score 39;  DB 22;  Length 751;
  Best Local Similarity   75.0%;  Pred. No. 1.5e+02;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 CLIQLFCK 9
              ||||||:|:
Db        486 CLIQLYCE 493


Search completed: October 14, 2003, 14:41:34
Job time : 40.7917 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 14:33:23 ; Search time 12.75 Seconds
(without alignments)
29.866 Million cell updates/sec

Title: US-09-881-490-191
Perfect score: 51
Sequence: 1 KCLIQLFCK 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 326717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	51	100.0	9	2	US-08-621-259A-191	Sequence 191, App
2	51	100.0	9	2	US-08-621-259A-192	Sequence 192, App
3	51	100.0	9	5	PCT-US95-09262-191	Sequence 191, App
4	51	100.0	9	5	PCT-US95-09262-192	Sequence 192, App
5	51	100.0	11	2	US-08-621-259A-193	Sequence 193, App
6	51	100.0	11	5	PCT-US95-09262-193	Sequence 193, App
7	36	70.6	78	3	US-08-905-223-452	Sequence 452, App
8	36	70.6	215	2	US-08-941-263-1	Sequence 1, Appli
9	36	70.6	215	3	US-09-227-178-1	Sequence 1, Appli
10	36	70.6	215	3	US-09-470-449-1	Sequence 1, Appli
11	36	70.6	215	4	US-09-726-775-1	Sequence 1, Appli
12	35	68.6	208	2	US-08-766-551-9	Sequence 9, Appli
13	35	68.6	259	4	US-09-075-454-1	Sequence 1, Appli
14	35	68.6	260	4	US-09-075-454-4	Sequence 4, Appli
15	34	66.7	223	3	US-08-820-970-2	Sequence 2, Appli
16	33.5	65.7	971	3	US-09-112-450-2	Sequence 2, Appli
17	33.5	65.7	971	4	US-09-419-291A-2	Sequence 2, Appli
18	33.5	65.7	2471	3	US-09-112-450-4	Sequence 4, Appli
19	33.5	65.7	2471	4	US-09-419-291A-4	Sequence 4, Appli
20	33	64.7	57	4	US-09-227-357-275	Sequence 275, App
21	33	64.7	493	3	US-09-090-808-4	Sequence 4, Appli
22	33	64.7	493	3	US-09-447-453-4	Sequence 4, Appli
23	33	64.7	735	4	US-09-585-858-10	Sequence 10, Appl
24	32	62.7	29	2	US-08-637-759B-304	Sequence 304, App
25	32	62.7	29	3	US-08-871-355A-304	Sequence 304, App
26	32	62.7	29	4	US-09-201-945-304	Sequence 304, App
27	32	62.7	200	4	US-09-198-452A-290	Sequence 290, App
28	32	62.7	221	4	US-09-511-881A-9	Sequence 9, Appli
29	32	62.7	344	3	US-09-120-365-63	Sequence 63, Appl
30	32	62.7	344	3	US-09-515-039-63	Sequence 63, Appl
31	32	62.7	420	2	US-08-164-292B-6	Sequence 6, Appli
32	32	62.7	420	3	US-08-845-623-6	Sequence 6, Appli
33	32	62.7	420	3	US-08-815-927-6	Sequence 6, Appli
34	32	62.7	420	4	US-09-103-330-6	Sequence 6, Appli
35	32	62.7	420	4	US-09-435-242-6	Sequence 6, Appli
36	32	62.7	503	4	US-09-562-737-68	Sequence 68, Appl
37	32	62.7	503	4	US-09-562-737-70	Sequence 70, Appl
38	32	62.7	815	4	US-09-914-259-18	Sequence 18, Appl
39	31	60.8	309	1	US-08-118-270-56	Sequence 56, Appl
40	31	60.8	309	5	PCT-US93-08528-56	Sequence 56, Appl
41	31	60.8	479	4	US-09-442-100-13	Sequence 13, Appl
42	31	60.8	479	4	US-08-939-106-13	Sequence 13, Appl
43	31	60.8	591	2	US-08-889-402-1	Sequence 1, Appli
44	31	60.8	605	2	US-08-889-402-2	Sequence 2, Appli
45	31	60.8	795	3	US-09-031-563-23	Sequence 23, Appl

ALIGNMENTS

```
RESULT 1
US-08-621-259A-191
; Sequence 191, Application US/08621259A
; Patent No. 5858974
;  GENERAL INFORMATION:
;    APPLICANT:  Little II, Roger G
;    APPLICANT:  Lim, Edward
;    APPLICANT:  Fadem, Mitchell B.
;    TITLE OF INVENTION:  Anti-Fungal Peptides
;    NUMBER OF SEQUENCES:  252
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McAndrews, Held & Malloy, Ltd.
;      STREET:  500 West Madison Street
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60661
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,259A
;      FILING DATE:  21-MAR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/504,841
;      FILING DATE:  20-JUL-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McNicholas, Janet M.
;      REGISTRATION NUMBER:  32,918
;      REFERENCE/DOCKET NUMBER:  11021US02
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/707-8889
;      TELEFAX:  312/707-9155
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  191:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.360"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1 & 9
;      OTHER INFORMATION:  /label= D-Lys
;      OTHER INFORMATION:  /note= "Positions 1 & 9 are D-lysine."
```

```
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  2
;      OTHER INFORMATION:  /label= D-Cys
;      OTHER INFORMATION:  /note= "Position 2 is D-cysteine."
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-08-621-259A-191

  Query Match           100.0%;  Score 51;  DB 2;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 2.5e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KCLIQLFCK 9
              |||||||||
Db          1 KCLIQLFCK 9


RESULT 2
US-08-621-259A-192
; Sequence 192, Application US/08621259A
; Patent No. 5858974
;  GENERAL INFORMATION:
;    APPLICANT:  Little II, Roger G
;    APPLICANT:  Lim, Edward
;    APPLICANT:  Fadem, Mitchell B.
;    TITLE OF INVENTION:  Anti-Fungal Peptides
;    NUMBER OF SEQUENCES:  252
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McAndrews, Held & Malloy, Ltd.
;      STREET:  500 West Madison Street
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60661
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,259A
;      FILING DATE:  21-MAR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/504,841
;      FILING DATE:  20-JUL-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McNicholas, Janet M.
;      REGISTRATION NUMBER:  32,918
;      REFERENCE/DOCKET NUMBER:  11021US02
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/707-8889
;      TELEFAX:  312/707-9155
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  192:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.361"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1 & 9
;      OTHER INFORMATION:  /label= D-Lys
;      OTHER INFORMATION:  /note= "Positions 1 & 9 are D-lysine."
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-08-621-259A-192

  Query Match           100.0%;  Score 51;  DB 2;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 2.5e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KCLIQLFCK 9
              |||||||||
Db          1 KCLIQLFCK 9


RESULT 3
PCT-US95-09262-191
; Sequence 191, Application PC/TUS9509262
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  Anti-Fungal Peptides
;    NUMBER OF SEQUENCES:  206
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/09262
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/372,105
;      FILING DATE:  13-JAN-95
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/306,473
;      FILING DATE:  15-SEP-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/273,540
;      FILING DATE:  11-JUL-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/209,762
;      FILING DATE:  11-MAR-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/183,222
;      FILING DATE:  14-JAN-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/093,202
;      FILING DATE:  15-JUL-93
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/030,644
;      FILING DATE:  12-MAR-93
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25,447
;      REFERENCE/DOCKET NUMBER:  27129/10040
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  191:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
```

```
;       OTHER INFORMATION:  "XMP.360"
;     FEATURE:
;       NAME/KEY:  Modified-site
;       LOCATION:  1 & 9
;       OTHER INFORMATION:  /label= D-Lys
;       OTHER INFORMATION:  /note= "Positions 1 & 9 are D-lysine."
;     FEATURE:
;       NAME/KEY:  Modified-site
;       LOCATION:  2
;       OTHER INFORMATION:  /label= D-Cys
;       OTHER INFORMATION:  /note= "Position 2 is D-cysteine."
;     FEATURE:
;       NAME/KEY:  Modified-site
;       LOCATION:  C-Terminus
;       OTHER INFORMATION:  /label= Amidation
;       OTHER INFORMATION:  /note= "The C-Terminus is Amidated"
PCT-US95-09262-191

  Query Match             100.0%;  Score 51;  DB 5;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.5e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KCLIQLFCK 9
              |||||||||
Db          1 KCLIQLFCK 9


RESULT 4
PCT-US95-09262-192
; Sequence 192, Application PC/TUS9509262
;  GENERAL INFORMATION:
;     APPLICANT:
;     TITLE OF INVENTION:  Anti-Fungal Peptides
;     NUMBER OF SEQUENCES:  206
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;       STREET:  6300 Sears Tower, 233 South Wacker Drive
;       CITY:  Chicago
;       STATE:  Illinois
;       COUNTRY:  United States of America
;       ZIP:  60606-6402
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US95/09262
;       FILING DATE:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/372,105
;       FILING DATE:  13-JAN-95
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/306,473
;       FILING DATE:  15-SEP-94
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/273,540
;       FILING DATE:  11-JUL-94
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/209,762
;       FILING DATE:  11-MAR-94
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/183,222
;       FILING DATE:  14-JAN-94
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/093,202
;       FILING DATE:  15-JUL-93
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/030,644
;       FILING DATE:  12-MAR-93
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Borun, Michael F.
;       REGISTRATION NUMBER:  25,447
;       REFERENCE/DOCKET NUMBER:  27129/10040
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  312/474-6300
;       TELEFAX:  312/474-0448
;       TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  192:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  9 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  circular
;     MOLECULE TYPE:  peptide
;     FEATURE:
;       NAME/KEY:  misc_feature
;       OTHER INFORMATION:  "XMP.361"
;     FEATURE:
;       NAME/KEY:  Modified-site
;       LOCATION:  1 & 9
;       OTHER INFORMATION:  /label= D-Lys
;       OTHER INFORMATION:  /note= "Positions 1 & 9 are D-lysine."
;     FEATURE:
;       NAME/KEY:  Modified-site
;       LOCATION:  C-Terminus
;       OTHER INFORMATION:  /label= Amidation
;       OTHER INFORMATION:  /note= "The C-Terminus is Amidated"
PCT-US95-09262-192

  Query Match             100.0%;  Score 51;  DB 5;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.5e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KCLIQLFCK 9
              |||||||||
Db          1 KCLIQLFCK 9


RESULT 5
US-08-621-259A-193
; Sequence 193, Application US/08621259A
; Patent No. 5858974
;  GENERAL INFORMATION:
;     APPLICANT:  Little II, Roger G
;     APPLICANT:  Lim, Edward
;     APPLICANT:  Fadem, Mitchell B.
;     TITLE OF INVENTION:  Anti-Fungal Peptides
;     NUMBER OF SEQUENCES:  252
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  McAndrews, Held & Malloy, Ltd.
;       STREET:  500 West Madison Street
;       CITY:  Chicago
;       STATE:  Illinois
;       COUNTRY:  United States of America
;       ZIP:  60661
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/621,259A
;       FILING DATE:  21-MAR-1996
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/504,841
;       FILING DATE:  20-JUL-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  McNicholas, Janet M.
;       REGISTRATION NUMBER:  32,918
;       REFERENCE/DOCKET NUMBER:  11021US02
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  312/707-8889
;       TELEFAX:  312/707-9155
;       TELEX:
;  INFORMATION FOR SEQ ID NO:  193:
;     SEQUENCE CHARACTERISTICS:
```

```
;       LENGTH:  11 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  circular
;     MOLECULE TYPE:  peptide
;     FEATURE:
;       NAME/KEY:  misc_feature
;       OTHER INFORMATION:  "XMP.362"
;     FEATURE:
;       NAME/KEY:  Modified-site
;       LOCATION:  1 & 11
;       OTHER INFORMATION:  /label= D-Lys
;       OTHER INFORMATION:  /note= "Positions 1 & 11 are D-lysine."
;     FEATURE:
;       NAME/KEY:  Modified-site
;       LOCATION:  C-Terminus
;       OTHER INFORMATION:  /label= Amidation
;       OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-08-621-259A-193

  Query Match             100.0%;  Score 51;  DB 2;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 0.0092;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KCLIQLFCK 9
              |||||||||
Db          2 KCLIQLFCK 10


RESULT 6
PCT-US95-09262-193
; Sequence 193, Application PC/TUS9509262
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  Anti-Fungal Peptides
;    NUMBER OF SEQUENCES:  206
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/09262
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/372,105
;      FILING DATE:  13-JAN-95
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/306,473
;      FILING DATE:  15-SEP-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/273,540
;      FILING DATE:  11-JUL-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/209,762
;      FILING DATE:  11-MAR-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/183,222
;      FILING DATE:  14-JAN-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/093,202
;      FILING DATE:  15-JUL-93
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/030,644
;      FILING DATE:  12-MAR-93
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25,447
;      REFERENCE/DOCKET NUMBER:  27129/10040
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  193:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  11 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.362"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1 & 11
;      OTHER INFORMATION:  /label= D-Lys
;      OTHER INFORMATION:  /note= "Positions 1 & 11 are D-lysine."
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated"
PCT-US95-09262-193

  Query Match             100.0%;  Score 51;  DB 5;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 0.0092;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KCLIQLFCK 9
              |||||||||
Db          2 KCLIQLFCK 10


RESULT 7
US-08-905-223-452
; Sequence 452, Application US/08905223
; Patent No. 6222029
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Jean-Baptiste D.
;    APPLICANT:  Duelert, Aymeric
;    APPLICANT:  Lacroix, Bruno
;    TITLE OF INVENTION:  5' ESTs FOR SECRETED PROTEINS
;    NUMBER OF SEQUENCES:  503
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Knobbe, Martens, Olson & Bear
;      STREET:  501 West Broadway
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92101-3505
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Win95
;      SOFTWARE:  Word
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/905,223
;      FILING DATE:
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Israelsen, Ned A.
;      REGISTRATION NUMBER:  29,655
;      REFERENCE/DOCKET NUMBER:
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 235-8550
;      TELEFAX:  (619) 235-0176
;  INFORMATION FOR SEQ ID NO:  452:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  78 amino acids
;      TYPE:  AMINO ACID
```

```
;      TOPOLOGY:  LINEAR
;    MOLECULE TYPE:  PROTEIN
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo Sapiens
;      TISSUE TYPE:  Brain
;    FEATURE:
;      NAME/KEY:  sig_peptide
;      LOCATION:  -69..-1
;      IDENTIFICATION METHOD:  Von Heijne matrix
;      OTHER INFORMATION:  score 5.7
;      OTHER INFORMATION:  seq LCRLLCLVRLFCC/SS
US-08-905-223-452

  Query Match             70.6%;  Score 36;  DB 3;  Length 78;
  Best Local Similarity   71.4%;  Pred. No. 18;
  Matches    5;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 CLIQLFC 8
              ||::|||
Db         62 CLVRLFC 68


RESULT 8
US-08-941-263-1
; Sequence 1, Application US/08941263
; Patent No. 5962231
;  GENERAL INFORMATION:
;    APPLICANT:  Yue, Henry
;    APPLICANT:  Corley, Neil C.
;    APPLICANT:  Shah, Purvi
;    TITLE OF INVENTION:  HUMAN GLUTATHIONE-S-TRANSFERASE
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Dr.
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/941,263
;      FILING DATE:  Filed Herewith
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0398 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  215 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  BRSTNOT16
;      CLONE:  3037341
US-08-941-263-1

  Query Match             70.6%;  Score 36;  DB 2;  Length 215;
  Best Local Similarity   66.7%;  Pred. No. 46;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KCLIQLFCK 9
                | |::||||
Db         86 KMLLELFCK 94


RESULT 9
US-09-227-178-1
; Sequence 1, Application US/09227178
; Patent No. 6071513
;  GENERAL INFORMATION:
;    APPLICANT:  Yue, Henry
;    APPLICANT:  Corley, Neil C.
;    APPLICANT:  Shah, Purvi
;    TITLE OF INVENTION:  HUMAN GLUTATHIONE-S-TRANSFERASE
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Dr.
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/227,178
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/941,263
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0398 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  215 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  BRSTNOT16
;      CLONE:  3037341
US-09-227-178-1

  Query Match             70.6%;  Score 36;  DB 3;  Length 215;
  Best Local Similarity   66.7%;  Pred. No. 46;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KCLIQLFCK 9
                | |::||||
Db         86 KMLLELFCK 94


RESULT 10
US-09-470-449-1
; Sequence 1, Application US/09470449
; Patent No. 6177546
;  GENERAL INFORMATION:
;    APPLICANT:  Yue, Henry
;    APPLICANT:  Corley, Neil C.
;    APPLICANT:  Shah, Purvi
;    TITLE OF INVENTION:  HUMAN GLUTATHIONE-S-TRANSFERASE
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Dr.
;      CITY:  Palo Alto
```

```
;       STATE:  CA
;       COUNTRY:  USA
;       ZIP:  94304
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette
;       COMPUTER:  IBM Compatible
;       OPERATING SYSTEM:  DOS
;       SOFTWARE:  FastSEQ for Windows Version 2.0
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/470,449
;       FILING DATE:  Herewith
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  09/227,178
;       FILING DATE:  January 8, 1999
;       APPLICATION NUMBER:  08/941,263
;       FILING DATE:  September 30, 1997
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Diana Hamlet-Cox
;       REGISTRATION NUMBER:  33,302
;       REFERENCE/DOCKET NUMBER:  PF-0398-2 DIV
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  650-855-0555
;       TELEFAX:  650-849-8886
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  215 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     IMMEDIATE SOURCE:
;       LIBRARY:  BRSTNOT16
;       CLONE:  3037341
US-09-470-449-1

  Query Match             70.6%;  Score 36;  DB 3;  Length 215;
  Best Local Similarity   66.7%;  Pred. No. 46;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KCLIQLFCK 9
              | |::||||
Db         86 KMLLELFCK 94


RESULT 11
US-09-726-775-1
; Sequence 1, Application US/09726775
; Patent No. 6506571
;   GENERAL INFORMATION:
;       APPLICANT: Yue, Henry
;                  Corley, Neil C.
;                  Shah, Purvi
;       TITLE OF INVENTION: HUMAN GLUTATHIONE-S-TRANSFERASE
;       NUMBER OF SEQUENCES: 3
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Incyte Pharmaceuticals, Inc.
;            STREET: 3174 Porter Dr.
;            CITY: Palo Alto
;            STATE: CA
;            COUNTRY: USA
;            ZIP: 94304
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Diskette
;            COMPUTER: IBM Compatible
;            OPERATING SYSTEM: DOS
;            SOFTWARE: FastSEQ for Windows Version 2.0
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/726,775
;            FILING DATE: 29-No. 6506571-2000
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/470,449
;            FILING DATE: <Unknown>
;            APPLICATION NUMBER: 08/941,263
;            FILING DATE: September 30, 1997
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Diana Hamlet-Cox
;            REGISTRATION NUMBER: 33,302
;            REFERENCE/DOCKET NUMBER: PF-0398-2 DIV
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 650-855-0555
;            TELEFAX: 650-849-8886
;   INFORMATION FOR SEQ ID NO: 1:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 215 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       IMMEDIATE SOURCE:
;            LIBRARY: BRSTNOT16
;            CLONE: 3037341
;       SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-726-775-1

  Query Match             70.6%;  Score 36;  DB 4;  Length 215;
  Best Local Similarity   66.7%;  Pred. No. 46;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KCLIQLFCK 9
              | |::||||
Db         86 KMLLELFCK 94


RESULT 12
US-08-766-551-9
; Sequence 9, Application US/08766551
; Patent No. 5840569
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Hawkins, Phillip R.
;    APPLICANT:  Goli, Surya K.
;    TITLE OF INVENTION:  NOVEL HUMAN GTP-BINDING PROTEINS
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  US
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/766,551
;      FILING DATE:  Herewith
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0168 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  208 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
```

```
;      LIBRARY:  GenBank
;      CLONE:  240986
US-08-766-551-9

  Query Match             68.6%;  Score 35;  DB 2;  Length 208;
  Best Local Similarity   75.0%;  Pred. No. 65;
  Matches    6;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 CLIQLFCK 9
              ||| |||
Db         16 CLINRFCK 23


RESULT 13
US-09-075-454-1
; Sequence 1, Application US/09075454
; Patent No. 6391580
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Tang, Y. Tom
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Guegler, Karl J.
;    APPLICANT:  Corley, Neil C.
;    APPLICANT:  Patterson, Chandra
;    APPLICANT:  Batra, Sajeev
;    APPLICANT:  Baughn, Mariah R.
;    TITLE OF INVENTION:  RAS PROTEINS
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  US
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  Word Perfect 6.1/MS-DOS 6.2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/075,454
;      FILING DATE:  Herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/766,551
;      FILING DATE:  DECEMBER 12, 1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Cerrone, Michael C.
;      REGISTRATION NUMBER:  39,132
;      REFERENCE/DOCKET NUMBER:  PF-0168-1 CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  259 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  KIDNNOT05
;      CLONE:  627565
US-09-075-454-1

  Query Match             68.6%;  Score 35;  DB 4;  Length 259;
  Best Local Similarity   75.0%;  Pred. No. 79;
  Matches    6;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 CLIQLFCK 9
              ||| |||
Db         67 CLINRFCK 74


RESULT 14
US-09-075-454-4
; Sequence 4, Application US/09075454
; Patent No. 6391580
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Tang, Y. Tom
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Guegler, Karl J.
;    APPLICANT:  Corley, Neil C.
;    APPLICANT:  Patterson, Chandra
;    APPLICANT:  Batra, Sajeev
;    APPLICANT:  Baughn, Mariah R.
;    TITLE OF INVENTION:  RAS PROTEINS
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  US
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  Word Perfect 6.1/MS-DOS 6.2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/075,454
;      FILING DATE:  Herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/766,551
;      FILING DATE:  DECEMBER 12, 1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Cerrone, Michael C.
;      REGISTRATION NUMBER:  39,132
;      REFERENCE/DOCKET NUMBER:  PF-0168-1 CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  260 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  PROSTUT08
;      CLONE:  1651593
US-09-075-454-4

  Query Match             68.6%;  Score 35;  DB 4;  Length 260;
  Best Local Similarity   75.0%;  Pred. No. 80;
  Matches    6;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 CLIQLFCK 9
              ||| |||
Db         68 CLINRFCK 75


RESULT 15
US-08-820-970-2
; Sequence 2, Application US/08820970
; Patent No. 6008022
;  GENERAL INFORMATION:
;    APPLICANT:  LI, YI
;    APPLICANT:  SU, KUI
;    APPLICANT:  LI, HAODONG
;    TITLE OF INVENTION:  HUMAN CYTOKINE POLYPEPTIDE
```

```
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HUMAN GENOME SCIENCES, INC.
;      STREET:  9410 KEY WEST AVENUE
;      CITY:  ROCKVILLE
;      STATE:  MARYLAND
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/820,970
;      FILING DATE:  20-MAR-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brookes, Anders, A.
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PF270
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  301-309-8504
;      TELEFAX:  301-309-8512
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  223 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-820-970-2

  Query Match             66.7%;  Score 34;  DB 3;  Length 223;
  Best Local Similarity   66.7%;  Pred. No. 1e+02;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 KCLIQLFCK 9
              |||||  |:
Db         37 KCLIQELCQ 45


Search completed: October 14, 2003, 14:42:47
Job time : 13.75 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 14:34:14 ; Search time 21.9375 Seconds
(without alignments)
66.104 Million cell updates/sec

Title: US-09-881-490-191
Perfect score: 51
Sequence: 1 KCLIQLFCK 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 600653 seqs, 161128416 residues

Total number of hits satisfying chosen parameters: 600653

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	51	100.0	9	9	US-09-881-490-191	Sequence 191, App
2	51	100.0	9	9	US-09-881-490-192	Sequence 192, App
3	51	100.0	11	9	US-09-881-490-193	Sequence 193, App
4	39	76.5	291	11	US-09-832-522-52	Sequence 52, Appl
5	39	76.5	297	11	US-09-832-522-34	Sequence 34, Appl
6	39	76.5	297	11	US-09-832-522-49	Sequence 49, Appl
7	39	76.5	298	11	US-09-832-522-44	Sequence 44, Appl
8	39	76.5	319	11	US-09-966-459A-11	Sequence 11, Appl
9	37	72.5	37	15	US-10-072-602B-172	Sequence 172, App
10	37	72.5	37	15	US-10-072-602B-558	Sequence 558, App
11	37	72.5	79	15	US-10-072-602B-171	Sequence 171, App
12	36	70.6	24	11	US-09-798-889-168	Sequence 168, App
13	36	70.6	195	11	US-09-798-889-164	Sequence 164, App
14	35	68.6	197	10	US-09-764-868-1119	Sequence 1119, Ap
15	35	68.6	228	10	US-09-764-868-1107	Sequence 1107, Ap
16	35	68.6	230	10	US-09-764-868-685	Sequence 685, App
17	35	68.6	259	14	US-10-051-986-1	Sequence 1, Appli
18	35	68.6	259	15	US-10-153-668-3	Sequence 3, Appli
19	35	68.6	260	14	US-10-051-986-4	Sequence 4, Appli
20	35	68.6	287	10	US-09-764-868-700	Sequence 700, App
21	35	68.6	734	15	US-10-170-528-2	Sequence 2, Appli
22	34	66.7	160	12	US-10-017-161-1380	Sequence 1380, Ap
23	34	66.7	222	9	US-09-897-772-2	Sequence 2, Appli
24	34	66.7	222	14	US-10-092-925-2	Sequence 2, Appli
25	34	66.7	223	10	US-09-811-088-8	Sequence 8, Appli
26	34	66.7	223	15	US-10-314-410-8	Sequence 8, Appli
27	34	66.7	318	10	US-09-886-055-101	Sequence 101, App
28	34	66.7	318	11	US-09-804-291-101	Sequence 101, App
29	34	66.7	318	15	US-10-081-775-5	Sequence 5, Appli
30	34	66.7	665	11	US-09-533-029-76	Sequence 76, Appl
31	34	66.7	665	16	US-10-278-536-100	Sequence 100, App
32	33.5	65.7	971	14	US-10-116-048-2	Sequence 2, Appli
33	33.5	65.7	2471	14	US-10-116-048-4	Sequence 4, Appli
34	33	64.7	43	11	US-09-764-891-3189	Sequence 3189, Ap
35	33	64.7	43	15	US-10-205-428-360	Sequence 360, App
36	33	64.7	57	11	US-09-983-802-275	Sequence 275, App
37	33	64.7	78	11	US-09-798-889-46	Sequence 46, Appl
38	33	64.7	87	9	US-09-945-301-7	Sequence 7, Appli
39	33	64.7	162	9	US-09-925-302-593	Sequence 593, App
40	33	64.7	163	11	US-09-898-659-2	Sequence 2, Appli
41	33	64.7	163	11	US-09-898-659-4	Sequence 4, Appli
42	33	64.7	240	10	US-09-925-300-1658	Sequence 1658, Ap
43	33	64.7	326	10	US-09-886-055-267	Sequence 267, App
44	33	64.7	326	11	US-09-804-291-267	Sequence 267, App
45	33	64.7	381	10	US-09-919-497-96	Sequence 96, Appl

ALIGNMENTS

RESULT 1
US-09-881-490-191
; Sequence 191, Application US/09881490
; Patent No. US20020077298A1
; GENERAL INFORMATION:
; APPLICANT: Little II, Roger G.
; Lim, Edward
; Fadem, Mitchell B.
; TITLE OF INVENTION: Anti-Fungal Peptides
; NUMBER OF SEQUENCES: 211
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.
; STREET: 500 West Madison Street, 34th FloorDrive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60661
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/881,490
; FILING DATE: 14-Jun-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/119,858
; FILING DATE: <Unknown>
; APPLICATION NUMBER: 08/372,105
; FILING DATE: 13-JAN-95
; APPLICATION NUMBER: 08/306,473
; FILING DATE: 15-SEP-94
; APPLICATION NUMBER: 08/273,540
; FILING DATE: 11-JUL-94
; APPLICATION NUMBER: 08/209,762
; FILING DATE: 11-MAR-94
; APPLICATION NUMBER: 08/183,222
; FILING DATE: 14-JAN-94

```
;              APPLICATION NUMBER: 08/093,202
;              FILING DATE: 15-JUL-93
;              APPLICATION NUMBER: 08/030,644
;              FILING DATE: 12-MAR-93
;         ATTORNEY/AGENT INFORMATION:
;              NAME: McNicholas, Janet M.
;              REGISTRATION NUMBER: 32,918
;              REFERENCE/DOCKET NUMBER: 100-238/11021US01
;         TELECOMMUNICATION INFORMATION:
;              TELEPHONE: 312/707-8889
;              TELEFAX: 312/707-9155
;              TELEX: 650 388-1248
;    INFORMATION FOR SEQ ID NO: 191:
;         SEQUENCE CHARACTERISTICS:
;              LENGTH: 9 amino acids
;              TYPE: amino acid
;              TOPOLOGY: circular
;         MOLECULE TYPE: peptide
;         FEATURE:
;              NAME/KEY:  misc_feature
;              OTHER INFORMATION:  "XMP.360"
;         FEATURE:
;              NAME/KEY:  Modified-site
;              LOCATION:  1 & 9
;              OTHER INFORMATION:  /label= D-Lys
;              /note= "Positions 1 & 9 are D-lysine."
;         FEATURE:
;              NAME/KEY:  Modified-site
;              LOCATION:  2
;              OTHER INFORMATION:  /label= D-Cys
;              /note= "Position 2 is D-cysteine."
;         FEATURE:
;              NAME/KEY:  Modified-site
;              LOCATION:  C-Terminus
;              OTHER INFORMATION:  /label= Amidation
;              /note= "The C-Terminus is Amidated"
;         SEQUENCE DESCRIPTION: SEQ ID NO: 191:
US-09-881-490-191

  Query Match             100.0%;  Score 51;  DB 9;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 5.4e-05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KCLIQLFCK 9
              |||||||||
Db          1 KCLIQLFCK 9


RESULT 2
US-09-881-490-192
; Sequence 192, Application US/09881490
; Patent No. US20020077298A1
;   GENERAL INFORMATION:
;         APPLICANT: Little II, Roger G.
;                    Lim, Edward
;                    Fadem, Mitchell B.
;         TITLE OF INVENTION: Anti-Fungal Peptides
;         NUMBER OF SEQUENCES: 211
;         CORRESPONDENCE ADDRESS:
;              ADDRESSEE: McAndrews, Held & Malloy, Ltd.
;              STREET: 500 West Madison Street, 34th FloorDrive
;              CITY: Chicago
;              STATE: Illinois
;              COUNTRY: United States of America
;              ZIP: 60661
;         COMPUTER READABLE FORM:
;              MEDIUM TYPE: Floppy disk
;              COMPUTER: IBM PC compatible
;              OPERATING SYSTEM: PC-DOS/MS-DOS
;              SOFTWARE: PatentIn Release #1.0, Version #1.25
;         CURRENT APPLICATION DATA:
;              APPLICATION NUMBER: US/09/881,490
;              FILING DATE: 14-Jun-2001
;         PRIOR APPLICATION DATA:
;              APPLICATION NUMBER: 09/119,858
;              FILING DATE: <Unknown>
;              APPLICATION NUMBER: 08/372,105
;              FILING DATE: 13-JAN-95
;              APPLICATION NUMBER: 08/306,473
;              FILING DATE: 15-SEP-94
;              APPLICATION NUMBER: 08/273,540
;              FILING DATE: 11-JUL-94
;              APPLICATION NUMBER: 08/209,762
;              FILING DATE: 11-MAR-94
;              APPLICATION NUMBER: 08/183,222
;              FILING DATE: 14-JAN-94
;              APPLICATION NUMBER: 08/093,202
;              FILING DATE: 15-JUL-93
;              APPLICATION NUMBER: 08/030,644
;              FILING DATE: 12-MAR-93
;         ATTORNEY/AGENT INFORMATION:
;              NAME: McNicholas, Janet M.
;              REGISTRATION NUMBER: 32,918
;              REFERENCE/DOCKET NUMBER: 100-238/11021US01
;         TELECOMMUNICATION INFORMATION:
;              TELEPHONE: 312/707-8889
;              TELEFAX: 312/707-9155
;              TELEX: 650 388-1248
;    INFORMATION FOR SEQ ID NO: 192:
;         SEQUENCE CHARACTERISTICS:
;              LENGTH: 9 amino acids
;              TYPE: amino acid
;              TOPOLOGY: circular
;         MOLECULE TYPE: peptide
;         FEATURE:
;              NAME/KEY:  misc_feature
;              OTHER INFORMATION:  "XMP.361"
;         FEATURE:
;              NAME/KEY:  Modified-site
;              LOCATION:  1 & 9
;              OTHER INFORMATION:  /label= D-Lys
;              /note= "Positions 1 & 9 are D-lysine."
;         FEATURE:
;              NAME/KEY:  Modified-site
;              LOCATION:  C-Terminus
;              OTHER INFORMATION:  /label= Amidation
;              /note= "The C-Terminus is Amidated"
;         SEQUENCE DESCRIPTION: SEQ ID NO: 192:
US-09-881-490-192

  Query Match             100.0%;  Score 51;  DB 9;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 5.4e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KCLIQLFCK 9
              |||||||||
Db          1 KCLIQLFCK 9


RESULT 3
US-09-881-490-193
; Sequence 193, Application US/09881490
; Patent No. US20020077298A1
;   GENERAL INFORMATION:
;         APPLICANT: Little II, Roger G.
;                    Lim, Edward
;                    Fadem, Mitchell B.
;         TITLE OF INVENTION: Anti-Fungal Peptides
;         NUMBER OF SEQUENCES: 211
;         CORRESPONDENCE ADDRESS:
;              ADDRESSEE: McAndrews, Held & Malloy, Ltd.
;              STREET: 500 West Madison Street, 34th FloorDrive
;              CITY: Chicago
;              STATE: Illinois
;              COUNTRY: United States of America
;              ZIP: 60661
```

```
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/881,490
;            FILING DATE: 14-Jun-2001
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/119,858
;            FILING DATE: <Unknown>
;            APPLICATION NUMBER: 08/372,105
;            FILING DATE: 13-JAN-95
;            APPLICATION NUMBER: 08/306,473
;            FILING DATE: 15-SEP-94
;            APPLICATION NUMBER: 08/273,540
;            FILING DATE: 11-JUL-94
;            APPLICATION NUMBER: 08/209,762
;            FILING DATE: 11-MAR-94
;            APPLICATION NUMBER: 08/183,222
;            FILING DATE: 14-JAN-94
;            APPLICATION NUMBER: 08/093,202
;            FILING DATE: 15-JUL-93
;            APPLICATION NUMBER: 08/030,644
;            FILING DATE: 12-MAR-93
;       ATTORNEY/AGENT INFORMATION:
;            NAME: McNicholas, Janet M.
;            REGISTRATION NUMBER: 32,918
;            REFERENCE/DOCKET NUMBER: 100-238/11021US01
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 312/707-8889
;            TELEFAX: 312/707-9155
;            TELEX: 650 388-1248
;   INFORMATION FOR SEQ ID NO: 193:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 11 amino acids
;            TYPE: amino acid
;            TOPOLOGY: circular
;       MOLECULE TYPE: peptide
;       FEATURE:
;            NAME/KEY:  misc_feature
;            OTHER INFORMATION:  "XMP.362"
;       FEATURE:
;            NAME/KEY:  Modified-site
;            LOCATION:  1 & 11
;            OTHER INFORMATION:  /label= D-Lys
;            /note= "Positions 1 & 11 are D-lysine."
;       FEATURE:
;            NAME/KEY:  Modified-site
;            LOCATION:  C-Terminus
;            OTHER INFORMATION:  /label= Amidation
;            /note= "The C-Terminus is Amidated"
;       SEQUENCE DESCRIPTION: SEQ ID NO: 193:
US-09-881-490-193

  Query Match             100.0%;  Score 51;  DB 9;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 0.031;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KCLIQLFCK 9
              |||||||||
Db          2 KCLIQLFCK 10


RESULT 4
US-09-832-522-52
; Sequence 52, Application US/09832522
; Publication No. US20030091563A1
; GENERAL INFORMATION:
;  APPLICANT: Shenoy, Suresh G
;  APPLICANT:  Gangolli, Esha A
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Smithson, Glennda
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Vernet, Corine
;  APPLICANT:  Wolenc, Adam R
;  APPLICANT:  Casman, Stacie J
;  APPLICANT:  Tchernev, Velizar T
;  APPLICANT:  Szekeres, Edward S
;  APPLICANT:  Grosse, William
;  APPLICANT:  Alsobrook, John P
;  APPLICANT:  Burgess, Catherine E
;  TITLE OF INVENTION: No. US20030091563A1el GPCR-Proteins and Nucleic Acids Encodi
;  FILE REFERENCE: 15966-767
;  CURRENT APPLICATION NUMBER: US/09/832,522
;  CURRENT FILING DATE:  2001-04-11
;  PRIOR APPLICATION NUMBER: 60/195,994
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: 60/196,538
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: 60/220,644
;  PRIOR FILING DATE: 2000-07-25
;  PRIOR APPLICATION NUMBER: 60/264,851
;  PRIOR FILING DATE: 2001-01-29
;  PRIOR APPLICATION NUMBER: 60/199,964
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: 60/268,567
;  PRIOR FILING DATE: 2001-02-14
;  PRIOR APPLICATION NUMBER: 60/199,955
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: 60/259,641
;  PRIOR FILING DATE: 2001-01-04
;  PRIOR APPLICATION NUMBER: 60/200,176
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: 60/199,948
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: 60/199,956
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: 60/218,995
;  PRIOR FILING DATE: 2000-07-17
;  NUMBER OF SEQ ID NOS: 112
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 52
;   LENGTH: 291
;   TYPE: PRT
;   ORGANISM: Gallus gallus
US-09-832-522-52

  Query Match             76.5%;  Score 39;  DB 11;  Length 291;
  Best Local Similarity   85.7%;  Pred. No. 51;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 CLIQLFC 8
              ||||:||
Db         79 CLIQMFC 85


RESULT 5
US-09-832-522-34
; Sequence 34, Application US/09832522
; Publication No. US20030091563A1
; GENERAL INFORMATION:
;  APPLICANT: Shenoy, Suresh G
;  APPLICANT:  Gangolli, Esha A
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Smithson, Glennda
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Vernet, Corine
;  APPLICANT:  Wolenc, Adam R
;  APPLICANT:  Casman, Stacie J
;  APPLICANT:  Tchernev, Velizar T
;  APPLICANT:  Szekeres, Edward S
;  APPLICANT:  Grosse, William
;  APPLICANT:  Alsobrook, John P
;  APPLICANT:  Burgess, Catherine E
;  TITLE OF INVENTION: No. US20030091563A1el GPCR-Proteins and Nucleic Acids Encodi
```

; FILE REFERENCE: 15966-767
; CURRENT APPLICATION NUMBER: US/09/832,522
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 60/195,994
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: 60/196,538
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: 60/220,644
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: 60/264,851
; PRIOR FILING DATE: 2001-01-29
; PRIOR APPLICATION NUMBER: 60/199,964
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/268,567
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/199,955
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/259,641
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/200,176
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/199,948
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/199,956
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/218,995
; PRIOR FILING DATE: 2000-07-17
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 34
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Gallus gallus
US-09-832-522-34

Query Match 76.5%; Score 39; DB 11; Length 297;
Best Local Similarity 85.7%; Pred. No. 52;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CLIQLFC 8
||||:||
Db 86 CLIQMFC 92

RESULT 6
US-09-832-522-49
; Sequence 49, Application US/09832522
; Publication No. US20030091563A1
; GENERAL INFORMATION:
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Gangolli, Esha A
; APPLICANT: Rastelli, Luca
; APPLICANT: Smithson, Glennda
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Vernet, Corine
; APPLICANT: Wolenc, Adam R
; APPLICANT: Casman, Stacie J
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Szekeres, Edward S
; APPLICANT: Grosse, William
; APPLICANT: Alsobrook, John P
; APPLICANT: Burgess, Catherine E
; TITLE OF INVENTION: No. US20030091563A1el GPCR-Proteins and Nucleic Acids Encoding Sa
; FILE REFERENCE: 15966-767
; CURRENT APPLICATION NUMBER: US/09/832,522
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 60/195,994
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: 60/196,538
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: 60/220,644
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: 60/264,851
; PRIOR FILING DATE: 2001-01-29
; PRIOR APPLICATION NUMBER: 60/199,964
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/268,567
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/199,955
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/259,641
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/200,176
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/199,948
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/199,956
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/218,995
; PRIOR FILING DATE: 2000-07-17
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 49
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Gallus gallus
US-09-832-522-49

Query Match 76.5%; Score 39; DB 11; Length 297;
Best Local Similarity 85.7%; Pred. No. 52;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CLIQLFC 8
||||:||
Db 86 CLIQMFC 92

RESULT 7
US-09-832-522-44
; Sequence 44, Application US/09832522
; Publication No. US20030091563A1
; GENERAL INFORMATION:
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Gangolli, Esha A
; APPLICANT: Rastelli, Luca
; APPLICANT: Smithson, Glennda
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Vernet, Corine
; APPLICANT: Wolenc, Adam R
; APPLICANT: Casman, Stacie J
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Szekeres, Edward S
; APPLICANT: Grosse, William
; APPLICANT: Alsobrook, John P
; APPLICANT: Burgess, Catherine E
; TITLE OF INVENTION: No. US20030091563A1el GPCR-Proteins and Nucleic Acids Encodin
; FILE REFERENCE: 15966-767
; CURRENT APPLICATION NUMBER: US/09/832,522
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 60/195,994
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: 60/196,538
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: 60/220,644
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: 60/264,851
; PRIOR FILING DATE: 2001-01-29
; PRIOR APPLICATION NUMBER: 60/199,964
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/268,567
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/199,955
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/259,641
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/200,176

```
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: 60/199,948
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: 60/199,956
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: 60/218,995
;  PRIOR FILING DATE: 2000-07-17
;  NUMBER OF SEQ ID NOS: 112
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 44
;   LENGTH: 298
;   TYPE: PRT
;   ORGANISM: Gallus gallus
US-09-832-522-44

  Query Match            76.5%;  Score 39;  DB 11;  Length 298;
  Best Local Similarity  85.7%;  Pred. No. 52;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 CLIQLFC 8
              ||||:||
Db         86 CLIQMFC 92


RESULT 8
US-09-966-459A-11
; Sequence 11, Application US/09966459A
; Publication No. US20030022237A1
; GENERAL INFORMATION:
;  APPLICANT: FEDER, J.N.
;  APPLICANT:  MINTIER, G.
;  APPLICANT:  RAMANATHAN, C.S.
;  APPLICANT:  HAWKEN, D.R.
;  APPLICANT:  CACACE, A.
;  APPLICANT:  BARBER, L.
;  APPLICANT:  KORNACKER, M.G.
;  TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPRBMY4,
;  TITLE OF INVENTION:  EXPRESSED HIGHLY IN PROSTATE, COLON, AND LUNG
;  FILE REFERENCE: D0039NP
;  CURRENT APPLICATION NUMBER: US/09/966,459A
;  CURRENT FILING DATE:  2001-09-26
;  PRIOR APPLICATION NUMBER: 60/235,833
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: 60/261,776
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: 60/305,351
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: 60/313,202
;  PRIOR FILING DATE: 2001-08-17
;  NUMBER OF SEQ ID NOS: 60
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
;   LENGTH: 319
;   TYPE: PRT
;   ORGANISM: CHICKEN
US-09-966-459A-11

  Query Match            76.5%;  Score 39;  DB 11;  Length 319;
  Best Local Similarity  85.7%;  Pred. No. 55;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 CLIQLFC 8
              ||||:||
Db         97 CLIQMFC 103


RESULT 9
US-10-072-602B-172
; Sequence 172, Application US/10072602B
; Publication No. US20030109670A1
; GENERAL INFORMATION:
;  APPLICANT: University of Utah Research Foundation
;  APPLICANT:  Cognetix, Inc.
;  APPLICANT:  Olivera, Baldomero M.
;  APPLICANT:  McIntosh, J, Michael
;  APPLICANT:  Watkins, Maren
;  APPLICANT:  Garrett, James E.
;  APPLICANT:  Cruz, Lourdes J.
;  APPLICANT:  Grilley, Michelle
;  APPLICANT:  Schoenfeld, Robert M.
;  APPLICANT:  Walker, Craig
;  APPLICANT:  Shetty, Reshma
;  APPLICANT:  Jones, Robert M.
;  TITLE OF INVENTION: Cone Snail Peptides
;  FILE REFERENCE: 2314-249
;  CURRENT APPLICATION NUMBER: US/10/072,602B
;  CURRENT FILING DATE:  2002-02-11
;  PRIOR APPLICATION NUMBER: US 60/267,408
;  PRIOR FILING DATE: 2001-02-09
;  NUMBER OF SEQ ID NOS: 638
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 172
;   LENGTH: 37
;   TYPE: PRT
;   ORGANISM: Conus purpurascens
;   FEATURE:
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)..(37)
;   OTHER INFORMATION: Xaa at residues 4, 7, 8, 11, 13 and 14 is Glu or gamma-carbc
;   OTHER INFORMATION: Glu; Xaa at residues 3 and 25 is Pro or hydroxy-Pro
US-10-072-602B-172

  Query Match            72.5%;  Score 37;  DB 15;  Length 37;
  Best Local Similarity  62.5%;  Pred. No. 18;
  Matches    5;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KCLIQLFC 8
              |||: :||
Db         30 KCLLSIFC 37


RESULT 10
US-10-072-602B-558
; Sequence 558, Application US/10072602B
; Publication No. US20030109670A1
; GENERAL INFORMATION:
;  APPLICANT: University of Utah Research Foundation
;  APPLICANT:  Cognetix, Inc.
;  APPLICANT:  Olivera, Baldomero M.
;  APPLICANT:  McIntosh, J, Michael
;  APPLICANT:  Watkins, Maren
;  APPLICANT:  Garrett, James E.
;  APPLICANT:  Cruz, Lourdes J.
;  APPLICANT:  Grilley, Michelle
;  APPLICANT:  Schoenfeld, Robert M.
;  APPLICANT:  Walker, Craig
;  APPLICANT:  Shetty, Reshma
;  APPLICANT:  Jones, Robert M.
;  TITLE OF INVENTION: Cone Snail Peptides
;  FILE REFERENCE: 2314-249
;  CURRENT APPLICATION NUMBER: US/10/072,602B
;  CURRENT FILING DATE:  2002-02-11
;  PRIOR APPLICATION NUMBER: US 60/267,408
;  PRIOR FILING DATE: 2001-02-09
;  NUMBER OF SEQ ID NOS: 638
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 558
;   LENGTH: 37
;   TYPE: PRT
;   ORGANISM: Conus purpurascens
US-10-072-602B-558

  Query Match            72.5%;  Score 37;  DB 15;  Length 37;
  Best Local Similarity  62.5%;  Pred. No. 18;
  Matches    5;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy          1 KCLIQLFC 8
              ||| : :|'
Db         30 KCLLSIFC 37
```

```
RESULT 11
US-10-072-602B-171
; Sequence 171, Application US/10072602B
; Publication No. US20030109670A1
; GENERAL INFORMATION:
;  APPLICANT: University of Utah Research Foundation
;  APPLICANT:  Cognetix, Inc.
;  APPLICANT:  Olivera, Baldomero M.
;  APPLICANT:  McIntosh, J, Michael
;  APPLICANT:  Watkins, Maren
;  APPLICANT:  Garrett, James E.
;  APPLICANT:  Cruz, Lourdes J.
;  APPLICANT:  Grilley, Michelle
;  APPLICANT:  Schoenfeld, Robert M.
;  APPLICANT:  Walker, Craig
;  APPLICANT:  Shetty, Reshma
;  APPLICANT:  Jones, Robert M.
;  TITLE OF INVENTION: Cone Snail Peptides
;  FILE REFERENCE: 2314-249
;  CURRENT APPLICATION NUMBER: US/10/072,602B
;  CURRENT FILING DATE:  2002-02-11
;  PRIOR APPLICATION NUMBER: US 60/267,408
;  PRIOR FILING DATE: 2001-02-09
;  NUMBER OF SEQ ID NOS: 638
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 171
;   LENGTH: 79
;   TYPE: PRT
;   ORGANISM: Conus purpurascens
US-10-072-602B-171

  Query Match           72.5%;  Score 37;  DB 15;  Length 79;
  Best Local Similarity  62.5%;  Pred. No. 34;
  Matches    5;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KCLIQLFC 8
              ||| : :||
Db         72 KCLLSIFC 79
```

```
RESULT 12
US-09-798-889-168
; Sequence 168, Application US/09798889
; Publication No. US20030004324A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: 31 Human secreted proteins
;  FILE REFERENCE: PZ026P1
;  CURRENT APPLICATION NUMBER: US/09/798,889
;  CURRENT FILING DATE:  2001-03-06
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/393,022
;  PRIOR FILING DATE: EARLIER FILING DATE: 1999-09-09
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/077,714
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-12
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/077,686
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-12
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/077,687
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-12
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/077,696
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-12
;  NUMBER OF SEQ ID NOS: 185
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 168
;   LENGTH: 24
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-798-889-168

  Query Match           70.6%;  Score 36;  DB 11;  Length 24;
  Best Local Similarity  66.7%;  Pred. No. 18;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KCLIQLFCK 9
              | |::||||
Db         14 KMLLELFCK 22
```

```
RESULT 13
US-09-798-889-164
; Sequence 164, Application US/09798889
; Publication No. US20030004324A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: 31 Human secreted proteins
;  FILE REFERENCE: PZ026P1
;  CURRENT APPLICATION NUMBER: US/09/798,889
;  CURRENT FILING DATE:  2001-03-06
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/393,022
;  PRIOR FILING DATE: EARLIER FILING DATE: 1999-09-09
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/077,714
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-12
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/077,686
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-12
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/077,687
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-12
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/077,696
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-12
;  NUMBER OF SEQ ID NOS: 185
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 164
;   LENGTH: 195
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (11)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-798-889-164

  Query Match           70.6%;  Score 36;  DB 11;  Length 195;
  Best Local Similarity  66.7%;  Pred. No. 1.1e+02;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KCLIQLFCK 9
              | |::||||
Db        103 KMLLELFCK 111
```

```
RESULT 14
US-09-764-868-1119
; Sequence 1119, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PTZ32
;  CURRENT APPLICATION NUMBER: US/09/764,868
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - refer to PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 1510
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1119
;   LENGTH: 197
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (187)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
;   LOCATION: (191)
```

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-868-1119

 Query Match 68.6%; Score 35; DB 10; Length 197;
 Best Local Similarity 75.0%; Pred. No. 1.6e-02;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 CLIQLFCK 9
 ||| |||
Db 95 CLINRFCK 102

RESULT 15
US-09-764-868-1107
; Sequence 1107, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ32
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1107
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (216)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-868-1107

 Query Match 68.6%; Score 35; DB 10; Length 228;
 Best Local Similarity 75.0%; Pred. No. 1.9e-02;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 CLIQLFCK 9
 ||| |||
Db 44 CLINRFCK 51

Search completed: October 14, 2003, 14:44:52
Job time : 22.9375 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 14:33:18 ; Search time 11.8125 Seconds
(without alignments)
73.271 Million cell updates/sec

Title: US-09-881-490-191
Perfect score: 51
Sequence: 1 KCLIQLFCK 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	38	74.5	313	2	E45774	odorant receptor 8
2	38	74.5	661	2	I52603	MPS1 protein - mou
3	37	72.5	163	2	F85046	hypothetical prote
4	37	72.5	259	2	PN0687	cyclin - anthracno
5	36	70.6	257	2	E82871	dimethyladenosine
6	35	68.6	208	2	A41636	GTP-binding protei
7	35	68.6	319	2	T44833	glycosyl transfera
8	35	68.6	341	2	T25051	hypothetical prote
9	35	68.6	432	2	AF0567	probable permease
10	35	68.6	462	2	G96772	hypothetical prote
11	35	68.6	584	2	E71643	poly-beta-hydroxyb
12	35	68.6	973	2	H86428	hypothetical prote
13	35	68.6	5376	2	T42215	zonadhesin - mouse
14	34	66.7	265	2	T14645	hypothetical prote
15	34	66.7	635	2	T00011	ccal protein - rat
16	34	66.7	665	2	D96621	auxin response fac
17	34	66.7	846	2	JC7720	acetyltransferase
18	34	66.7	1142	2	T46155	hypothetical prote
19	34	66.7	1289	2	E90098	RNA polymerase III
20	34	66.7	2747	2	B49132	fat facets (faf) s
21	33.5	65.7	2471	2	T03820	probable histidine
22	33	64.7	126	2	T23246	hypothetical prote
23	33	64.7	127	2	E81134	hypothetical prote
24	33	64.7	175	2	T23245	hypothetical prote
25	33	64.7	340	2	T32104	hypothetical prote
26	33	64.7	381	1	A47327	selenoprotein P pr
27	33	64.7	504	2	T48992	hypothetical prote
28	33	64.7	522	2	A31556	glucose transport
29	33	64.7	523	2	S06920	glucose transport
30	33	64.7	577	2	T09024	proline-rich prote
31	33	64.7	645	2	F86627	LysR family transc
32	33	64.7	727	2	T25362	hypothetical prote
33	33	64.7	942	2	S49109	spherulin - Melolo
34	33	64.7	1084	1	S19661	DNA-directed DNA p
35	33	64.7	1086	2	T43266	DNA-directed DNA p
36	33	64.7	1320	2	S57113	BUD4 protein - yea
37	32.5	63.7	505	2	T28172	hypothetical prote
38	32.5	63.7	666	2	T03090	hypothetical prote
39	32	62.7	77	2	B82473	hypothetical prote
40	32	62.7	160	2	E71197	hypothetical prote
41	32	62.7	169	2	G72096	hypothetical prote
42	32	62.7	169	2	E86525	hypothetical prote
43	32	62.7	232	2	T20934	hypothetical prote
44	32	62.7	312	2	A86279	F14L17.21 protein
45	32	62.7	328	2	S73842	thymidylate syntha

ALIGNMENTS

RESULT 1
E45774
odorant receptor 8 - channel catfish
C;Species: Ictalurus punctatus (channel catfish)
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999
C;Accession: E45774
R;Ngai, J.; Dowling, M.M.; Buck, L.; Axel, R.; Chess, A.
Cell 72, 657-666, 1993
A;Title: The family of genes encoding odorant receptors in the channel catfish.
A;Reference number: A45774; MUID:93201590; PMID:7916654
A;Accession: E45774
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-313 <NGA>
A;Experimental source: olfactory epithelium
A;Note: sequence extracted from NCBI backbone (NCBIP:127748)
C;Superfamily: olfactory receptor OR14
C;Keywords: olfaction; transmembrane protein

Query Match 74.5%; Score 38; DB 2; Length 313;
Best Local Similarity 71.4%; Pred. No. 10;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 CLIQLFC 8
              ||:|:||
Db         98 CLVQMFC 104
```

RESULT 2
I52603
MPS1 protein - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 05-Nov-1999
C;Accession: I52603
R;Spilsbury, K.; O'Mara, M.A.; Wu, W.M.; Rowe, P.B.; Symonds, G.; Takayama, Y.
Blood 85, 1620-1629, 1995
A;Title: Isolation of a novel macrophage-specific gene by differential cDNA analysis
A;Reference number: I52603; MUID:95195232; PMID:7888681
A;Accession: I52603
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-661 <RES>
A;Cross-references: GB:L20315; NID:g431419; PIDN:AAA73957.1; PID:g431420

Query Match 74.5%; Score 38; DB 2; Length 661;
Best Local Similarity 55.6%; Pred. No. 19;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 KCLIQLFCK 9
              || :::|||
Db        434 KCTLKIFCK 442
```

```
RESULT 3
F85046
hypothetical protein AT4g03680 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
C;Accession: F85046
R;anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sprin
Nature 402, 769-777, 1999
A;Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A;Reference number: A85001; MUID:20083488; PMID:10617198
A;Accession: F85046
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-163 <STO>
A;Cross-references: GB:NC_001268; NID:g7270691; PIDN:CAB77853.1; GSPDB:GN00140
C;Genetics:
A;Gene: AT4g03680
A;Map position: 4

  Query Match            72.5%;  Score 37;  DB 2;  Length 163;
  Best Local Similarity  100.0%;  Pred. No. 9.4;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KCLIQLF 7
              |||||||
Db          7 KCLIQLF 13


RESULT 4
PN0687
cyclin - anthracnose fungus (Colletotrichum gloeosporioides) (fragment)
C;Species: Glomerella cingulata, Colletotrichum gloeosporioides
C;Date: 05-Aug-1994 #sequence_revision 05-Aug-1994 #text_change 17-Mar-1999
C;Accession: PN0687
R;Masel, A.M.; Struijk, N.; McIntyre, C.L.; Irwin, J.A.G.; Manners, J.M.
Gene 133, 141-145, 1993
A;Title: A strain-specific cyclin homolog in the fungal phytopathogen Colletotrichum glo
A;Reference number: PN0687; MUID:94040789; PMID:8224886
A;Accession: PN0687
A;Molecule type: mRNA
A;Residues: 1-259 <MAS>
C;Superfamily: cyclin
F;48-193/Region: cyclin box

  Query Match            72.5%;  Score 37;  DB 2;  Length 259;
  Best Local Similarity  85.7%;  Pred. No. 14;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 CLIQLFC 8
              | ||||| 
Db        186 CFIQLFC 192


RESULT 5
E82871
dimethyladenosine transferase UU601 [imported] - Ureaplasma urealyticum
C;Species: Ureaplasma urealyticum
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000
C;Accession: E82871
R;Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.
submitted to GenBank, February 2000
A;Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a min
A;Reference number: A82870
A;Accession: E82871
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-257 <GLA>
A;Cross-references: GB:AE002158; GB:AF222894; NID:g6899599; PIDN:AAF31015.1; GSPDB:GN001
A;Experimental source: serovar 3; biovar 1
C;Genetics:
A;Gene: ksgA; UU601
A;Genetic code: SGC3
C;Superfamily: rRNA (adenine-N6-)-methyltransferase

  Query Match            70.6%;  Score 36;  DB 2;  Length 257;
  Best Local Similarity  85.7%;  Pred. No. 21;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LIQLFCK 9
              |:|||||
Db        152 LVQLFCK 158


RESULT 6
A41636
GTP-binding protein rah - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 05-Jun-1992 #sequence_revision 05-Jun-1992 #text_change 02-Feb-2001
C;Accession: A41636
R;Morimoto, B.H.; Chuang, C.C.; Koshland Jr., D.E.
Genes Dev. 5, 2386-2391, 1991
A;Title: Molecular cloning of a member of a new class of low-molecular-weight GTP-bi
A;Reference number: A41636; MUID:92090719; PMID:1752434
A;Accession: A41636
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-208 <MOR>
A;Cross-references: GB:S72304; NID:g240985; PIDN:AAB20669.1; PID:g240986
C;Superfamily: ras transforming protein; translation elongation factor Tu homology
C;Keywords: membrane trafficking; nucleotide binding; P-loop
F;8-15/Region: nucleotide-binding motif A (P-loop)

  Query Match            68.6%;  Score 35;  DB 2;  Length 208;
  Best Local Similarity  75.0%;  Pred. No. 27;
  Matches    6;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 CLIQLFCK 9
              |||  |||
Db         16 CLINRFCK 23


RESULT 7
T44833
glycosyl transferase homolog [imported] - Acinetobacter lwoffii
C;Species: Acinetobacter lwoffii
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 19-May-2000
C;Accession: T44833
R;Nakar, D.; Gutnick, D.L.
submitted to the EMBL Data Library, July 1999
A;Description: Genomic organization of the wce region of Acinetobacter lwoffii RAG-1
A;Reference number: Z22856
A;Accession: T44833
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-319 <NAK>
A;Cross-references: EMBL:AJ243431; PIDN:CAB57201.1
A;Experimental source: strain RAG-1
C;Genetics:
A;Gene: weeD
C;Superfamily: Acinetobacter lwoffii glycosyl transferase homolog

  Query Match            68.6%;  Score 35;  DB 2;  Length 319;
  Best Local Similarity  62.5%;  Pred. No. 38;
  Matches    5;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 CLIQLFCK 9
              ||| :||:
Db        197 CLIDIFCE 204


RESULT 8
T25051
hypothetical protein T21B4.6 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
```

```
C;Accession: T25051
R;Smye, R.
submitted to the EMBL Data Library, October 1996
A;Reference number: Z19974
A;Accession: T25051
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-341 <WIL>
A;Cross-references: EMBL:Z81124; PIDN:CAB03372.1; GSPDB:GN00020; CESP:T21B4.6
A;Experimental source: clone T21B4
C;Genetics:
A;Gene: CESP:T21B4.6
A;Map position: 2
A;Introns: 120/1; 183/3
C;Superfamily: Caenorhabditis hypothetical protein C49G7.2

  Query Match             68.6%;  Score 35;  DB 2;  Length 341;
  Best Local Similarity   77.8%;  Pred. No. 40;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 KCLIQLFCK 9
              | || ||||
Db        316 KFLISLFCK 324


RESULT 9
AF0567
probable permease protein STY0572 [imported] - Salmonella enterica subsp. enterica serov
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AF0567
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AF0567
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-432 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD05008.1; PID:g16501791; GSPDB:GN00176
C;Genetics:
A;Gene: STY0572

  Query Match             68.6%;  Score 35;  DB 2;  Length 432;
  Best Local Similarity   71.4%;  Pred. No. 49;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 CLIQLFC 8
              ||:| ||
Db         57 CLVQAFC 63


RESULT 10
G96772
hypothetical protein F1M20.8 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: G96772
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: G96772
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-462 <STO>
A;Cross-references: GB:AE005173; NID:g6539242; PIDN:AAF15912.1; GSPDB:GN00141
C;Genetics:
A;Gene: F1M20.8
A;Map position: 1

  Query Match             68.6%;  Score 35;  DB 2;  Length 462;
  Best Local Similarity   50.0%;  Pred. No. 52;
  Matches    4;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 CLIQLFCK 9
              |:: |||:
Db        317 CMVDLFCR 324


RESULT 11
E71643
poly-beta-hydroxybutyrate polymerase (phbC2) RP820 - Rickettsia prowazekii
C;Species: Rickettsia prowazekii
C;Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 03-Nov-2000
C;Accession: E71643
R;Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.; Alsmar
Nature 396, 133-140, 1998
A;Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria
A;Reference number: A71630; MUID:99039499; PMID:9823893
A;Accession: E71643
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-584 <AND>
A;Cross-references: GB:AJ235273; GB:AJ235269; NID:g3861237; PIDN:CAA15245.1; PID:g38
A;Experimental source: strain Madrid E
C;Genetics:
A;Gene: phbC2; RP820
C;Superfamily: poly(3-hydroxyalkanoic acid) synthase phbC

  Query Match             68.6%;  Score 35;  DB 2;  Length 584;
  Best Local Similarity   62.5%;  Pred. No. 62;
  Matches    5;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 CLIQLFCK 9
              |:|: |||
Db         51 CIIEQFCK 58


RESULT 12
H86428
hypothetical protein F26G16.3 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001
C;Accession: H86428
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Al
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Ki
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marz
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tal
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: H86428
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-973 <STO>
A;Cross-references: GB:AE005172; NID:g6634764; PIDN:AAF19744.1; GSPDB:GN00141
C;Genetics:
A;Map position: 1

  Query Match             68.6%;  Score 35;  DB 2;  Length 973;
```

```
  Best Local Similarity   85.7%;  Pred. No. 94;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 CLIQLFC 8
              | |||||
Db        810 CKIQLFC 816


RESULT 13
T42215
zonadhesin - mouse
N;Alternate names: sperm-specific membrane protein
C;Species: Mus musculus (house mouse)
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C;Accession: T42215
R;Gao, Z.; Garbers, D.L.
J. Biol. Chem. 273, 3415-3421, 1998
A;Title: Species diversity in the structure of zonadhesin, a sperm-specific membrane prc
A;Reference number: Z22080; MUID:98123114; PMID:9452463
A;Accession: T42215
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-5376 <GAO>
A;Cross-references: EMBL:U97068; NID:g3327420; PID:g3327421; PIDN:AAC26680.1
C;Genetics:
A;Gene: Zan
A;Map position: 5
C;Function:
A;Description: functions in multiple cell adhesion processes
A;Note: found exclusively on the apical region of the sperm head
C;Keywords: cell adhesion

  Query Match             68.6%;  Score 35;  DB 2;  Length 5376;
  Best Local Similarity   50.0%;  Pred. No. 3.8e+02;
  Matches    4;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KCLIQLFC 8
              ||::::||
Db       3231 KCVLRIFC 3238


RESULT 14
T14645
hypothetical protein 265 - Sorghum mitochondrion
C;Species: mitochondrion Sorghum bicolor (sorghum)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C;Accession: T14645
R;Tang, H.V.; Pring, D.R.; Muza, F.R.; Yan, B.
Curr. Genet. 29, 265-274, 1996
A;Title: Sorghum mitochondrial orf25 and a related chimeric configuration of a male-ster
A;Reference number: S65767; MUID:96163056; PMID:8595673
A;Accession: T14645
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-265 <VAN>
A;Cross-references: EMBL:U22068; NID:g733079; PIDN:AAA97555.1; PID:g733080
A;Experimental source: strain IS1112C; coleoptile
C;Genetics:
A;Genome: mitochondrion
C;Keywords: mitochondrion

  Query Match             66.7%;  Score 34;  DB 2;  Length 265;
  Best Local Similarity   71.4%;  Pred. No. 51;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 CLIQLFC 8
              ||: |||
Db        228 CLLMLFC 234


RESULT 15
T00011
ccal protein - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 21-Jul-2000
C;Accession: T00011
R;Hayashi, Y.; Kiyono, T.; Fujita, M.; Ishibashi, M.
J. Biol. Chem. 272, 18082-18086, 1997
A;Title: Ccal is required for formation of growth-arrested confluent monolayer of ra
A;Reference number: Z14036; MUID:97364729; PMID:9218439
A;Accession: T00011
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-635 <HAY>
A;Cross-references: EMBL:AB000215; NID:g2564000; PIDN:BAA22932.1; PID:g2564001
C;Genetics:
A;Gene: ccal
C;Superfamily: rat ccal protein

  Query Match             66.7%;  Score 34;  DB 2;  Length 635;
  Best Local Similarity   71.4%;  Pred. No. 1e+02;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 CLIQLFC 8
              | :||||
Db         53 CYLQLFC 59


Search completed: October 14, 2003, 14:38:08
Job time : 13.8125 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 14:33:18 ; Search time 6.375 Seconds
(without alignments)
66.391 Million cell updates/sec

Title: US-09-881-490-191
Perfect score: 51
Sequence: 1 KCLIQLFCK 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	39	76.5	886	1	VP39_HUMAN	Q96jc1 homo sapien
2	39	76.5	886	1	VP39_MOUSE	Q8r513 mus musculu
3	37	72.5	245	1	CG1_COLGL	Q06374 colletotric
4	36	70.6	277	1	KSGA_UREPA	Q9ppn8 ureaplasma
5	35	68.6	259	1	RB34_HUMAN	Q9bzq1 homo sapien
6	35	68.6	259	1	RB34_MOUSE	Q64008 mus musculu
7	35	68.6	5376	1	ZAN_MOUSE	O88799 mus musculu
8	34	66.7	318	1	OYD1_HUMAN	Q9h346 homo sapien
9	34	66.7	2778	1	FAF_DROME	P55824 drosophila
10	33	64.7	291	1	RT02_MOUSE	Q924t2 mus musculu
11	33	64.7	315	1	OXM1_HUMAN	Q9h3q1 homo sapien
12	33	64.7	381	1	SELP_HUMAN	P49908 homo sapien
13	33	64.7	522	1	GTR2_RAT	P12336 rattus norv
14	33	64.7	523	1	GTR2_MOUSE	P14246 mus musculu
15	33	64.7	602	1	S21F_MACFA	Q9gmu6 macaca fasc
16	33	64.7	712	1	S21F_HUMAN	Q9nyb5 homo sapien
17	33	64.7	1447	1	BUD4_YEAST	P47136 saccharomyc
18	32	62.7	416	1	TC10_HUMAN	Q12799 homo sapien
19	32	62.7	497	1	NU2M_CHOCR	P48903 *chondrus cr*
20	32	62.7	732	1	YLH3_SCHPO	Q9hgq3 schizosacch
21	32	62.7	815	1	KINH_CAEEL	P34540 caenorhabdi
22	32	62.7	947	1	PMA2_ARATH	P19456 arabidopsis
23	32	62.7	948	1	PMA1_ARATH	P20649 arabidopsis
24	32	62.7	960	1	CAP2_SORBI	P29194 sorghum bic
25	32	62.7	967	1	CAP2_MAIZE	P51059 zea mays (m
26	31	60.8	127	1	ATPZ_PROMO	P29705 propionigen
27	31	60.8	246	1	ENV_RSVSR	P03397 rous sarcom
28	31	60.8	279	1	LACR_STAXY	O33813 staphylococ
29	31	60.8	380	1	SELP_MOUSE	P70274 mus musculu
30	31	60.8	385	1	SELP_RAT	P25236 rattus norv
31	31	60.8	391	1	P53_RABIT	Q95330 oryctolagus
32	31	60.8	438	1	RUMA_HAEIN	P44643 haemophilus
33	31	60.8	480	1	YB9Q_YEAST	P38348 saccharomyc
34	31	60.8	503	1	C303_DROME	Q9v399 drosophila
35	31	60.8	506	1	MURE_LEPIN	Q8f4j4 leptospira
36	31	60.8	559	1	PAXI_CHICK	P49024 gallus gall
37	31	60.8	591	1	PAXI_HUMAN	P49023 homo sapien
38	31	60.8	993	1	DPOL_BPAPS	Q9t1q3 bacteriopha
39	31	60.8	998	1	GTFI_HUMAN	P78347 *h general t*
40	31	60.8	998	1	GTFI_MOUSE	Q9esz8 mus musculu
41	31	60.8	1318	1	YAB2_SCHPO	Q09804 schizosacch
42	31	60.8	2670	1	YAQ5_SCHPO	Q10105 schizosacch
43	30.5	59.8	685	1	CFAH_BOVIN	Q28085 *bos taurus*
44	30	58.8	74	1	SSPS_STRPY	P29840 streptococc
45	30	58.8	189	1	APOD_HUMAN	P05090 homo sapien

ALIGNMENTS

```
RESULT 1
VP39_HUMAN
ID   VP39_HUMAN     STANDARD;      PRT;   886 AA.
AC   Q96JC1; O94869; Q96B93; Q96RM0;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Vam6/Vps39-like protein (hVam6p).
GN   VPS39 OR KIAA0770.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC   TISSUE=Brain;
RX   MEDLINE=21342080; PubMed=11448994;
RA   Caplan S., Hartnell L.M., Aguilar R.C., Naslavsky N., Bonifacino J.S.;
RT   "Human Vam6p promotes lysosome clustering and fusion in vivo.";
RL   J. Cell Biol. 154:109-122(2001).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RC   TISSUE=Brain;
RX   MEDLINE=99087487; PubMed=9872452;
RA   Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Miyajima N.,
RA   Tanaka A., Kotani H., Nomura N., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes. XI.
RT   The complete sequences of 100 new cDNA clones from brain which code
RT   for large proteins in vitro.";
RL   DNA Res. 5:277-286(1998).
RN   [3]
RP   SEQUENCE OF 98-886 FROM N.A.
RC   TISSUE=Lung;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: May play a role in clustering and fusion of late
CC       endosomes and lysosomes.
```

```
CC   -!- SUBUNIT: Homooligomer.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic and associated with the
CC       membranes of lysosomes and late endosomes.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=Q96JC1-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=Q96JC1-2; Sequence=VSP_004075;
CC   -!- TISSUE SPECIFICITY: Ubiquitous. Highly expressed in heart,
CC       skeletal muscle, kidney, pancreas, brain, placenta and lung.
CC   -!- SIMILARITY: Contains 1 CNH domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF280814; AAK72222.1; -.
DR   EMBL; AF334400; AAK58862.1; -.
DR   EMBL; AB018313; BAA34490.2; ALT_INIT.
DR   EMBL; BC015817; AAH15817.1; ALT_INIT.
DR   InterPro; IPR001180; Citron.
DR   Pfam; PF00780; CNH; 1.
KW   Transport; Protein transport; Alternative splicing.
FT   DOMAIN       15    294       CNH.
FT   VARSPLIC     47     58       VPADVASPESGS -> G (in isoform 2).
FT                                /FTId=VSP_004075.
FT   CONFLICT     98    100       YVH -> GTR (IN REF. 3).
FT   CONFLICT    135    135       Q -> R (IN REF. 1; AF280814 AND 2).
SQ   SEQUENCE   886 AA;  101780 MW;  2C0C48753AFB6A7F CRC64;

  Query Match             76.5%;  Score 39;  DB 1;  Length 886;
  Best Local Similarity   75.0%;  Pred. No. 8.9;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 CLIQLFCK 9
              |||||:|:
Db        621 CLIQLYCE 628


RESULT 2
VP39_MOUSE
ID   VP39_MOUSE     STANDARD;      PRT;   886 AA.
AC   Q8R5L3; Q922I3;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Vam6/Vps39-like protein (Vps39 protein).
GN   VPS39 OR PLDN.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC   STRAIN=Swiss Webster / NIH;
RA   Falcon-Perez J.M., Dell'Angelica E.C.;
RL   Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RC   TISSUE=Breast tumor;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: May play a role in clustering and fusion of late
CC       endosomes and lysosomes (By similarity).
CC   -!- SUBUNIT: Homooligomer (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic and associated with the
CC       membranes of lysosomes and late endosomes (By similarity).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=Q8R5L3-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=Q8R5L3-2; Sequence=VSP_004076;
CC   -!- SIMILARITY: Contains 1 CNH domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF281050; AAL79766.1; -.
DR   EMBL; AF281051; AAL79767.1; -.
DR   EMBL; BC007479; AAH07479.1; -.
DR   MGD; MGI:1927580; Pldn.
DR   GO; GO:0005768; C:endosome; IDA.
DR   GO; GO:0006944; P:membrane fusion; IPI.
DR   InterPro; IPR001180; Citron.
DR   Pfam; PF00780; CNH; 1.
KW   Transport; Protein transport; Alternative splicing.
FT   DOMAIN       15    294       CNH.
FT   VARSPLIC     47     58       VPADVASPESGS -> G (in isoform 2).
FT                                /FTId=VSP_004076.
SQ   SEQUENCE   886 AA;  101692 MW;  BD46EF52D404C2DA CRC64;

  Query Match             76.5%;  Score 39;  DB 1;  Length 886;
  Best Local Similarity   75.0%;  Pred. No. 8.9;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 CLIQLFCK 9
              |||||:|:
Db        621 CLIQLYCE 628


RESULT 3
CG1_COLGL
ID   CG1_COLGL      STANDARD;      PRT;   245 AA.
AC   Q06374;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   01-FEB-1995 (Rel. 31, Last annotation update)
DE   Putative G1/S-specific cyclin.
GN   CYC1.
OS   Colletotrichum gloeosporioides (Anthracnose fungus) (Glomerella
OS   cingulata).
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Sordariomycetes incertae sedis; Phyllachorales; Phyllachoraceae;
OC   Glomerella.
OX   NCBI_TaxID=5457;
RN   [1]
```

```
RP   SEQUENCE FROM N.A.
RC   STRAIN=Biotype B / Isolate race-3;
RX   MEDLINE=94040789; PubMed=8224888;
RA   Masel A.M., Struijk N., McIntyre C.L., Irwin J.A.G., Manners J.M.;
RT   "A strain-specific cyclin homolog in the fungal phytopathogen
RT   Colletotrichum gloeosporioides.";
RL   Gene 133:141-145(1993).
CC   -!- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S
CC       (START) TRANSITION. INTERACTS WITH THE CDC2 PROTEIN KINASE TO
CC       FORM MPF (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. STRONGEST TO OTHER
CC       G1/S CYCLINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X69099; CAA48845.1; -.
DR   InterPro; IPR006670; Cyclin.
DR   InterPro; IPR006671; Cyclin_N.
DR   Pfam; PF00134; cyclin; 1.
DR   SMART; SM00385; CYCLIN; 1.
DR   PROSITE; PS00292; CYCLINS; 1.
KW   Cyclin; Cell cycle; Cell division.
SQ   SEQUENCE   245 AA;  28143 MW;  687FE836383F5B90 CRC64;

  Query Match             72.5%;  Score 37;  DB 1;  Length 245;
  Best Local Similarity   85.7%;  Pred. No. 5.9;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 CLIQLFC 8
              | ||:||
Db        172 CFIQLFC 178


RESULT 4
KSGA_UREPA
ID   KSGA_UREPA     STANDARD;      PRT;   277 AA.
AC   Q9PPN8;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Dimethyladenosine transferase (EC 2.1.1.-) (S-adenosylmethionine-6-N',
DE   N'-adenosyl(rRNA) dimethyltransferase) (16S rRNA dimethylase) (High
DE   level kasugamycin resistance protein ksgA) (Kasugamycin
DE   dimethyltransferase).
GN   KSGA OR UU601.
OS   Ureaplasma parvum (Ureaplasma urealyticum biotype 1).
OC   Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Ureaplasma.
OX   NCBI_TaxID=134821;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Serovar 3;
RX   MEDLINE=20500219; PubMed=11048724;
RA   Glass J.I., Lefkowitz E.J., Glass J.S., Heiner C.R., Chen E.Y.,
RA   Cassell G.H.;
RT   "The complete sequence of the mucosal pathogen Ureaplasma
RT   urealyticum.";
RL   Nature 407:757-762(2000).
CC   -!- FUNCTION: Specifically dimethylates two adjacent adenosines in the
CC       loop of a conserved hairpin near the 3'end of 16S rRNA in the 30S
CC       particle. Its inactivation leads to kasugamycin resistance (By
CC       similarity).
CC   -!- SIMILARITY: BELONGS TO THE RRNA ADENINE N-6-METHYLTRANSFERASE
CC       FAMILY. KSGA SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE002158; AAF31015.1; ALT_INIT.
DR   HSSP; P13956; 1QAM.
DR   HAMAP; MF_00607; -; 1.
DR   InterPro; IPR001737; RRNA_A_dimeth.
DR   InterPro; IPR000051; SAM_bind.
DR   Pfam; PF00398; RrnaAD; 1.
DR   SMART; SM00650; rADc; 1.
DR   TIGRFAMs; TIGR00755; ksgA; 1.
DR   PROSITE; PS01131; RRNA_A_DIMETH; 1.
KW   mRNA processing; Transferase; Methyltransferase;
KW   Antibiotic resistance; Complete proteome.
SQ   SEQUENCE   277 AA;  32145 MW;  EC4E13D6B1549BCE CRC64;

  Query Match             70.6%;  Score 36;  DB 1;  Length 277;
  Best Local Similarity   85.7%;  Pred. No. 10;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LIQLFCK 9
              |:|||||
Db        172 LVQLFCK 178


RESULT 5
RB34_HUMAN
ID   RB34_HUMAN     STANDARD;      PRT;   259 AA.
AC   Q9BZG1; Q8NCJ8;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Ras-related protein Rab-34 (Rab-39) (Ras-related protein Rah).
GN   RAB34 OR RAB39 OR RAH.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Hong W.;
RT   "Human Rab39 coding region (cDNA).";
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Skeletal muscle;
RA   Stanchi F., Lanfranchi G.;
RT   "Full-length sequencing of 100 cDNA clones from human adult skeletal
RT   muscle.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo, and Mammary gland;
RA   Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA   Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA   Wagatsuma M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA   Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA   Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA   Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
RA   Ninomiya K., Iwayanagi T.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: PROTEIN TRANSPORT. PROBABLY INVOLVED IN VESICULAR
CC       TRAFFIC (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE SMALL GTPASE SUPERFAMILY. RAB FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
```

```
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF322067; AAK09397.1; -.
DR   EMBL; AJ277106; CAC81760.1; -.
DR   EMBL; AK027312; BAB55034.1; -.
DR   EMBL; AK074689; BAC11141.1; -.
DR   HSSP; P17080; 1A2K.
DR   Genew; HGNC:16519; RAB34.
DR   InterPro; IPR003579; GTPase_Rab.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00175; RAB; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
KW   GTP-binding; Lipoprotein; Prenylation; Transport; Protein transport;
KW   Golgi stack; Polymorphism.
FT   NP_BIND      59     66       GTP (BY SIMILARITY).
FT   NP_BIND     107    111       GTP (BY SIMILARITY).
FT   NP_BIND     166    169       GTP (BY SIMILARITY).
FT   DOMAIN       81     89       EFFECTOR REGION (BY SIMILARITY).
FT   LIPID       257    257       GERANYL-GERANYL (POTENTIAL).
FT   LIPID       258    258       GERANYL-GERANYL (POTENTIAL).
FT   VARIANT     197    197       V -> L (in dbSNP:12125).
FT                                /FTId=VAR_015097.
FT   CONFLICT     55     55       V -> I (IN REF. 3; BAC11141).
SQ   SEQUENCE   259 AA;  29044 MW;  6D38A6F090F9802D CRC64;

  Query Match             68.6%;  Score 35;  DB 1;  Length 259;
  Best Local Similarity   75.0%;  Pred. No. 15;
  Matches    6;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 CLIQLFCK 9
              |||  |||
Db         67 CLINRFCK 74


RESULT 6
RB34_MOUSE
ID   RB34_MOUSE     STANDARD;      PRT;   259 AA.
AC   Q64008; Q99P59; Q99P90;
DT   30-MAY-2000 (Rel. 39, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Ras-related protein Rab-34 (Rab-39) (Ras-related protein Rah) (Ras-
DE   related homolog).
GN   RAB34 OR RAB39 OR RAH1 OR RAH.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J;
RA   Hong W.;
RT   "Mouse Rab39 coding region (cDNA).";
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Neuron;
RA   Hari M., Morimoto B.H., Asai D.J.;
RT   "Full length sequence of Rah, a novel member of the Rab family of
RT   small GTPases.";
RL   Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 52-259 FROM N.A.
RC   TISSUE=Neuron;
RX   MEDLINE=92090719; PubMed=1752434;
RA   Morimoto B.H., Chuang C.-C., Koshland D.E. Jr.;
RT   "Molecular cloning of a member of a new class of low-molecular-weight
RT   GTP-binding proteins.";
RL   Genes Dev. 5:2386-2391(1991).
CC   -!- FUNCTION: PROTEIN TRANSPORT. PROBABLY INVOLVED IN VESICULAR
CC       TRAFFIC (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE SMALL GTPASE SUPERFAMILY. RAB FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF322068; AAK09398.1; -.
DR   EMBL; AF327929; AAK11241.1; -.
DR   EMBL; S72304; AAB20669.1; -.
DR   PIR; A41636; A41636.
DR   HSSP; P06749; 1A2B.
DR   MGD; MGI:104606; Rab34.
DR   InterPro; IPR003579; GTPase_Rab.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00175; RAB; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
KW   GTP-binding; Lipoprotein; Prenylation; Transport; Protein transport;
KW   Golgi stack.
FT   NP_BIND      59     66       GTP (BY SIMILARITY).
FT   NP_BIND     107    111       GTP (BY SIMILARITY).
FT   NP_BIND     166    169       GTP (BY SIMILARITY).
FT   DOMAIN       81     89       EFFECTOR REGION (BY SIMILARITY).
FT   LIPID       257    257       GERANYL-GERANYL (POTENTIAL).
FT   LIPID       258    258       GERANYL-GERANYL (POTENTIAL).
FT   CONFLICT     52     52       I -> N (IN REF. 3).
SQ   SEQUENCE   259 AA;  29100 MW;  158F5132F0DDE808 CRC64;

  Query Match             68.6%;  Score 35;  DB 1;  Length 259;
  Best Local Similarity   75.0%;  Pred. No. 15;
  Matches    6;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 CLIQLFCK 9
              |||  |||
Db         67 CLINRFCK 74


RESULT 7
ZAN_MOUSE
ID   ZAN_MOUSE      STANDARD;      PRT;  5376 AA.
AC   O88799; O08647;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Zonadhesin precursor.
GN   ZAN.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Testis;
RX   MEDLINE=98123114; PubMed=9452463;
RA   Gao Z., Garbers D.L.;
RT   "Species diversity in the structure of zonadhesin, a sperm-specific
RT   membrane protein containing multiple cell adhesion molecule-like
RT   domains.";
RL   J. Biol. Chem. 273:3415-3421(1998).
RN   [2]
RP   SEQUENCE OF 4864-5376 FROM N.A.
RC   TISSUE=Testis;
RX   MEDLINE=97271566; PubMed=9126492;
RA   Gao Z., Harumi T., Garbers D.L.;
RT   "Chromosome localization of the mouse zonadhesin gene and the human
```

```
RT   zonadhesin gene (ZAN).";
RL   Genomics 41:119-122(1997).
CC   -!- FUNCTION: BINDS IN A SPECIES-SPECIFIC MANNER TO THE ZONA PELLUCIDA
CC       OF THE EGG. MAY BE INVOLVED IN GAMETE RECOGNITION AND/OR
CC       SIGNALING.
CC   -!- SUBUNIT: PROBABLY FORMS COVALENT OLIGOMERS.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN, EXCLUSIVELY ON THE
CC       APICAL REGION OF THE SPERM HEAD.
CC   -!- TISSUE SPECIFICITY: IN TESTIS; PRIMARILY IN HAPLOID SPERMATIDS.
CC   -!- DOMAIN: THE MAM DOMAINS PROBABLY MEDIATES SPERM ADHESION TO THE
CC       ZONA PELLUCIDA.
CC   -!- DOMAIN: DURING SPERM MIGRATION THROUGH THE REPRODUCTIVE TRACTS,
CC       THE MUCIN-LIKE DOMAIN MIGHT INHIBIT INAPPROPRIATE TRAPPING OF
CC       SPERMATOZOA OR PROMOTING ADHESION TO THE OVIDUCTAL ISTHMUS.
CC   -!- DOMAIN: THE VWFD DOMAIN 2 MAY MEDIATE COVALENT
CC       OLIGOMERIZATION (BY SIMILARITY TO HUMAN INTESTINAL MUCIN MUC2).
CC   -!- SIMILARITY: Contains 3 MAM domains.
CC   -!- SIMILARITY: Contains 25 VWFD domains.
CC   -!- SIMILARITY: Contains 1 EGF-like domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U97068; AAC26680.1; -.
DR   EMBL; U83190; AAC53125.1; -.
DR   PIR; T42215; T42215.
DR   MGD; MGI:106656; Zan.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR003645; FolN.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR000998; MAM_domain.
DR   InterPro; IPR002919; TIL_Cysrich.
DR   InterPro; IPR003328; TILa_Cysrich.
DR   InterPro; IPR001007; VWF_C.
DR   InterPro; IPR001846; VWF_D.
DR   Pfam; PF00629; MAM; 3.
DR   Pfam; PF01826; TIL; 25.
DR   Pfam; PF02345; TILa; 25.
DR   Pfam; PF00094; vwd; 4.
DR   SMART; SM00181; EGF; 2.
DR   SMART; SM00274; FOLN; 11.
DR   SMART; SM00137; MAM; 2.
DR   SMART; SM00214; VWC; 17.
DR   SMART; SM00216; VWD; 4.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 18.
DR   PROSITE; PS00740; MAM_1; FALSE_NEG.
DR   PROSITE; PS50060; MAM_2; 3.
KW   Signal; Glycoprotein; Transmembrane; EGF-like domain; Cell adhesion;
KW   Repeat.
FT   SIGNAL        1     17       POTENTIAL.
FT   CHAIN        18   5376       ZONADHESIN.
FT   DOMAIN       18   5310       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   5311   5337       POTENTIAL.
FT   DOMAIN     5338   5376       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       45    210       MAM 1.
FT   DOMAIN      215    374       MAM 2.
FT   DOMAIN      377    542       MAM 3.
FT   DOMAIN      547   1170       80 X HEPTAPEPTIDE REPEATS (APPROXIMATE)
FT                                (MUCIN-LIKE DOMAIN).
FT   DOMAIN     1171   1280       VWFD 1 (PARTIAL).
FT   DOMAIN     1281   1669       VWFD 2.
FT   DOMAIN     1670   2056       VWFD 3.
FT   DOMAIN     2057   2459       VWFD 4.
FT   DOMAIN     2460   2579       VWFD 5 (PARTIAL).
FT   DOMAIN     2580   2699       VWFD 6 (PARTIAL).
FT   DOMAIN     2700   2819       VWFD 7 (PARTIAL).
FT   DOMAIN     2820   2939       VWFD 8 (PARTIAL).
FT   DOMAIN     2940   3059       VWFD 9 (PARTIAL).
FT   DOMAIN     3060   3179       VWFD 10 (PARTIAL).
FT   DOMAIN     3180   3299       VWFD 11 (PARTIAL).
FT   DOMAIN     3300   3416       VWFD 12 (PARTIAL).
FT   DOMAIN     3417   3536       VWFD 13 (PARTIAL).
FT   DOMAIN     3537   3656       VWFD 14 (PARTIAL).
FT   DOMAIN     3657   3776       VWFD 15 (PARTIAL).
FT   DOMAIN     3777   3892       VWFD 16 (PARTIAL).
FT   DOMAIN     3893   4928       VWFD 17 (PARTIAL).
FT   DOMAIN     4029   4148       VWFD 18 (PARTIAL).
FT   DOMAIN     4149   4263       VWFD 19 (PARTIAL).
FT   DOMAIN     4264   4283       VWFD 20 (PARTIAL).
FT   DOMAIN     4384   4503       VWFD 21 (PARTIAL).
FT   DOMAIN     4504   4623       VWFD 22 (PARTIAL).
FT   DOMAIN     4624   4743       VWFD 23 (PARTIAL).
FT   DOMAIN     4744   4863       VWFD 24 (PARTIAL).
FT   DOMAIN     4864   5261       VWFD 25.
FT   DOMAIN     5259   5295       EGF-LIKE.
FT   DISULFID   5263   5274       BY SIMILARITY.
FT   DISULFID   5268   5283       BY SIMILARITY.
FT   DISULFID   5285   5294       BY SIMILARITY.
FT   CARBOHYD    339    339       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    499    499       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1216   1216       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1239   1239       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1314   1314       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1814   1814       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1908   1908       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1933   1933       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2028   2028       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2111   2111       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2142   2142       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2332   2332       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2533   2533       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2575   2575       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2692   2692       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2812   2812       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   3052   3052       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   3065   3065       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   3144   3144       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   3172   3172       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   3288   3288       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   3292   3292       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   3782   3782       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   4005   4005       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   4136   4136       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   4243   4243       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   4254   4254       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   4335   4335       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   4376   4376       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   4586   4586       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   5136   5136       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   5252   5252       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   5376 AA;  579908 MW;  0E44DB77DF2A2620 CRC64;

  Query Match           68.6%;  Score 35;  DB 1;  Length 5376;
  Best Local Similarity  50.0%;  Pred. No. 2.8e+02;
  Matches    4;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KCLIQLFC 8
              ||::::||
Db       3231 KCVLRIFC 3238


RESULT 8
OYD1_HUMAN
ID   OYD1_HUMAN     STANDARD;      PRT;   318 AA.
AC   Q9H346;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Olfactory receptor 52D1 (HOR5'beta14).
GN   OR52D1.
```

```
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20570519; PubMed=11121057;
RA   Bulger M., Bender M.A., van Doorninck J.H., Wertman B., Farrell C.M.,
RA   Felsenfeld G., Groudine M., Hardison R.;
RT   "Comparative structural and functional analysis of the olfactory
RT   receptor genes flanking the human and mouse beta-globin gene
RT   clusters.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:14560-14565(2000).
CC   -!- FUNCTION: PUTATIVE ODORANT RECEPTOR.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF137396; AAG41676.1; -.
DR   GO; GO:0016021; C:integral to membrane; NAS.
DR   GO; GO:0004984; F:olfactory receptor activity; NAS.
DR   GO; GO:0007608; P:olfaction; NAS.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Multigene family; Olfaction.
FT   DOMAIN        1     28       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     29     52       1 (POTENTIAL).
FT   DOMAIN       53     60       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     61     82       2 (POTENTIAL).
FT   DOMAIN       83    103       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    104    123       3 (POTENTIAL).
FT   DOMAIN      124    142       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    143    161       4 (POTENTIAL).
FT   DOMAIN      162    198       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    199    222       5 (POTENTIAL).
FT   DOMAIN      223    239       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    240    262       6 (POTENTIAL).
FT   DOMAIN      263    278       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    279    298       7 (POTENTIAL).
FT   DOMAIN      299    318       CYTOPLASMIC (POTENTIAL).
FT   DISULFID    100    192       BY SIMILARITY.
FT   CARBOHYD      5      5       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   318 AA;  35121 MW;  178968087D613CB2 CRC64;

  Query Match             66.7%;  Score 34;  DB 1;  Length 318;
  Best Local Similarity   71.4%;  Pred. No. 27;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 CLIQLFC 8
              || |:||
Db        100 CLAQMFC 106


RESULT 9
FAF_DROME
ID   FAF_DROME      STANDARD;      PRT;  2778 AA.
AC   P55824; Q9V9T6; Q9Y0Z7;
DT   01-NOV-1997 (Rel. 35, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Probable ubiquitin carboxyl-terminal hydrolase FAF (EC 3.1.2.15)
DE   (Ubiquitin thiolesterase FAF) (Ubiquitin-specific processing protease
DE   FAF) (Deubiquitinating enzyme FAF) (Fat facets protein).
GN   FAF OR BCDNA:LD22582 OR CG1945.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS 2 AND 3), AND TISSUE SPECIFICITY.
RC   TISSUE=Eye imaginal disk;
RX   MEDLINE=93202020; PubMed=1295747;
RA   Fischer-Vize J.A., Rubin G.M., Lehmann R.;
RT   "The fat facets gene is required for Drosophila eye and embryo
RT   development.";
RL   Development 116:985-1000(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [3]
RP   REVISIONS, AND ALTERNATIVE SPLICING.
RC   STRAIN=Berkeley;
RX   MEDLINE=22426069; PubMed=12537572;
RA   Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA   Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA   Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA   Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA   Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA   Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA   Lewis S.E.;
RT   "Annotation of the Drosophila melanogaster euchromatic genome: a
RT   systematic review.";
RL   Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN   [4]
RP   SEQUENCE OF 1089-2778 FROM N.A. (ISOFORM 1).
```

```
RC   STRAIN=Berkeley;
RX   MEDLINE=20196012; PubMed=10731138;
RA   Rubin G.M., Hong L., Brokstein P., Evans-Holm M., Frise E.,
RA   Stapleton M., Harvey D.A.;
RT   "A Drosophila complementary DNA resource.";
RL   Science 287:2222-2224(2000).
CC   -!- FUNCTION: REQUIRED FOR EYE AND EMBRYO DEVELOPMENT, AND PLAYS A
CC       ROLE IN COMPOUND EYE ASSEMBLY AND OOGENESIS RESPECTIVELY. IN THE
CC       LARVAL EYE DISKS, CELLS OUTSIDE THE ASSEMBLING FACETS REQUIRE THIS
CC       PROTEIN FOR SHORT-RANGE CELL INTERACTIONS THAT PREVENT THE MYSTERY
CC       CELLS FROM BECOMING PHOTORECEPTORS. IT IS ALSO REQUIRED FOR
CC       NUCLEAR MIGRATION AND CELLULARIZATION IN EARLY EMBRYOGENESIS AND
CC       COULD PLAY A ROLE IN POLE CELL DETERMINATION, DEVELOPMENT OR
CC       FUNCTION.
CC   -!- CATALYTIC ACTIVITY: Ubiquitin C-terminal thiolester + H(2)O =
CC       ubiquitin + a thiol.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=3;
CC         Comment=Experimental confirmation may be lacking for some
CC         isoforms;
CC       Name=1;
CC         IsoId=P55824-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=P55824-2; Sequence=VSP_005270;
CC       Name=3;
CC         IsoId=P55824-3; Sequence=VSP_005269;
CC   -!- TISSUE SPECIFICITY: EYE DISKS AND OVARIES.
CC   -!- DEVELOPMENTAL STAGE: Expressed both maternally and zygotically.
CC   -!- SIMILARITY: Belongs to peptidase family C19.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L04959; AAF01345.1; -.
DR   EMBL; L04958; AAF01346.1; -.
DR   EMBL; L04960; AAF01347.1; -.
DR   EMBL; L04960; AAF01348.1; -.
DR   EMBL; AE003779; AAF57198.1; -.
DR   EMBL; AE003779; AAN14291.1; -.
DR   EMBL; AF145677; AAD38652.1; -.
DR   MEROPS; C19.007; -.
DR   FlyBase; FBgn0005632; faf.
DR   GO; GO:0005737; C:cytoplasm; IDA.
DR   GO; GO:0007349; P:cellularization; IMP.
DR   GO; GO:0009795; P:embryonic morphogenesis; IMP.
DR   GO; GO:0008583; P:mystery cell fate differentiation (sensu Dr. . .; IMP.
DR   GO; GO:0007097; P:nuclear migration; IMP.
DR   GO; GO:0006512; P:ubiquitin cycle; IGI.
DR   InterPro; IPR001394; UCH-2.
DR   Pfam; PF00443; UCH; 1.
DR   PROSITE; PS00972; UCH_2_1; 1.
DR   PROSITE; PS00973; UCH_2_2; 1.
DR   PROSITE; PS50235; UCH_2_3; 1.
KW   Ubl conjugation pathway; Hydrolase; Thiol protease;
KW   Developmental protein; Vision; Alternative splicing.
FT   ACT_SITE   1677   1677       BY SIMILARITY.
FT   ACT_SITE   1978   1978       BY SIMILARITY.
FT   ACT_SITE   1986   1986       BY SIMILARITY.
FT   VARSPLIC   2705   2778       KCRRVIIKKLVESKDEEDATTATTAATTEVTISPATAIATA
FT                                ATLEPAGMSELTTMVEKNLIISQENPQAKSSLQ -> VTRA
FT                                NNV (in isoform 3).
FT                                /FTId=VSP_005269.
FT   VARSPLIC   2742   2778       IATAATLEPAGMSELTTMVEKNLIISQENPQAKSSLQ ->
FT                                SQRQQL (in isoform 2).
FT                                /FTId=VSP_005270.
FT   CONFLICT    234    234       E -> D (IN REF. 1).
FT   CONFLICT   2725   2725       T -> S (IN REF. 1; AAF01345).
SQ   SEQUENCE   2778 AA;  311139 MW;  FFB90438BA53A02B CRC64;
```

```
  Query Match             66.7%;  Score 34;  DB 1;  Length 2778;
  Best Local Similarity   77.8%;  Pred. No. 2.3e+02;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 KCLIQLFCK 9
              ||| ||| |
Db       2480 KCLTQLFHK 2488


RESULT 10
RT02_MOUSE
ID   RT02_MOUSE     STANDARD;      PRT;   291 AA.
AC   Q924T2;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Mitochondrial 28S ribosomal protein S2 (MRP-S2).
GN   MRPS2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21413863; PubMed=11402041;
RA   Suzuki T., Terasaki M., Takemoto-Hori C., Hanada T., Ueda T., Wada A.,
RA   Watanabe K.;
RT   "Proteomic analysis of the mammalian mitochondrial ribosome.
RT   Identification of protein components in the 28 S small subunit.";
RL   J. Biol. Chem. 276:33181-33195(2001).
CC   -!- SUBUNIT: Component of the mitochondrial ribosome small subunit
CC       (28S) which comprises a 12S rRNA and about 30 distinct proteins.
CC   -!- SUBCELLULAR LOCATION: Mitochondrial.
CC   -!- SIMILARITY: BELONGS TO THE S2P FAMILY OF RIBOSOMAL PROTEINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB055389; BAB62530.1; -.
DR   MGD; MGI:2153089; Mrps2.
DR   InterPro; IPR001865; Ribosomal_S2.
DR   InterPro; IPR005706; S2_bact_org.
DR   Pfam; PF00318; Ribosomal_S2; 1.
DR   PRINTS; PR00395; RIBOSOMALS2.
DR   TIGRFAMs; TIGR01011; rpsB_bact; 1.
DR   PROSITE; PS00962; RIBOSOMAL_S2_1; 1.
KW   Ribosomal protein; Mitochondrion.
SQ   SEQUENCE   291 AA;  32313 MW;  33CFE4731BCEBADD CRC64;

  Query Match             64.7%;  Score 33;  DB 1;  Length 291;
  Best Local Similarity   100.0%;  Pred. No. 38;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 IQLFCK 9
              ||||||
Db        243 IQLFCK 248


RESULT 11
OXM1_HUMAN
ID   OXM1_HUMAN     STANDARD;      PRT;   315 AA.
AC   Q9H341;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Olfactory receptor 51M1 (HOR5'beta7).
GN   OR51M1.
```

```
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20570519; PubMed=11121057;
RA   Bulger M., Bender M.A., van Doorninck J.H., Wertman B., Farrell C.M.,
RA   Felsenfeld G., Groudine M., Hardison R.;
RT   "Comparative structural and functional analysis of the olfactory
RT   receptor genes flanking the human and mouse beta-globin gene
RT   clusters.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:14560-14565(2000).
CC   -!- FUNCTION: PUTATIVE ODORANT RECEPTOR.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF137396; AAG41681.1; ALT_INIT.
DR   GO; GO:0016021; C:integral to membrane; NAS.
DR   GO; GO:0004984; F:olfactory receptor activity; NAS.
DR   GO; GO:0007608; P:olfaction; NAS.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Multigene family; Olfaction.
FT   DOMAIN        1     26       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     27     50       1 (POTENTIAL).
FT   DOMAIN       51     58       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     59     80       2 (POTENTIAL).
FT   DOMAIN       81    101       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    102    121       3 (POTENTIAL).
FT   DOMAIN      122    140       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    141    159       4 (POTENTIAL).
FT   DOMAIN      160    196       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    197    220       5 (POTENTIAL).
FT   DOMAIN      221    237       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    238    260       6 (POTENTIAL).
FT   DOMAIN      261    275       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    276    295       7 (POTENTIAL).
FT   DOMAIN      296    315       CYTOPLASMIC (POTENTIAL).
FT   DISULFID     98    190       BY SIMILARITY.
FT   CARBOHYD      5      5       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   315 AA;  35462 MW;  DC19D2F5456D2005 CRC64;

  Query Match           64.7%;  Score 33;  DB 1;  Length 315;
  Best Local Similarity  71.4%;  Pred. No. 41;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 CLIQLFC 8
              | ||:||
Db         98 CQIQMFC 104


RESULT 12
SELP_HUMAN
ID   SELP_HUMAN     STANDARD;      PRT:   381 AA.
AC   P49908;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Selenoprotein P precursor (SeP).
GN   SEPP1 OR SELP.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Heart, and Liver;
RX   MEDLINE=93133823; PubMed=8421687;
RA   Hill K.E., Lloyd R.S., Burk R.F.;
RT   "Conserved nucleotide sequences in the open reading frame and 3'
RT   untranslated region of selenoprotein P mRNA.";
RL   Proc. Natl. Acad. Sci. U.S.A. 90:537-541(1993).
RN   [2]
RP   PARTIAL SEQUENCE.
RX   MEDLINE=94191007; PubMed=8142465;
RA   Aakesson B., Bellew T., Burk R.F.;
RT   "Purification of selenoprotein P from human plasma.";
RL   Biochim. Biophys. Acta 1204:243-249(1994).
RN   [3]
RP   CHARACTERIZATION.
RX   MEDLINE=98413836; PubMed=9735174;
RA   Mostert V., Lombeck I., Abel J.;
RT   "A novel method for the purification of selenoprotein P from human
RT   plasma.";
RL   Arch. Biochem. Biophys. 357:326-330(1998).
RN   [4]
RP   CHARACTERIZATION.
RX   MEDLINE=20239644; PubMed=10775431;
RA   Mostert V.;
RT   "Selenoprotein P: properties, functions, and regulation.";
RL   Arch. Biochem. Biophys. 376:433-438(2000).
RN   [5]
RP   REVIEW.
RX   MEDLINE=95017128; PubMed=7931697;
RA   Burk R.F., Hill K.E.;
RT   "Selenoprotein P. A selenium-rich extracellular glycoprotein.";
RL   J. Nutr. 124:1891-1897(1994).
CC   -!- FUNCTION: MIGHT BE RESPONSIBLE FOR SOME OF THE EXTRACELLULAR
CC       ANTIOXIDANT DEFENSE PROPERTIES OF SELENIUM OR MIGHT BE INVOLVED IN
CC       THE TRANSPORT OF SELENIUM.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: MADE IN THE LIVER AND HEART AND SECRETED INTO
CC       THE PLASMA. IT IS ALSO FOUND IN THE KIDNEY.
CC   -!- MISCELLANEOUS: THE SELENOCYSTEINES ARE ALL ENCODED BY THE OPAL
CC       CODON, UGA.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z11793; CAA77836.1; -.
DR   PIR; A47327; A47327.
DR   Genew; HGNC:10751; SEPP1.
DR   MIM; 601484; -.
DR   GO; GO:0008430; F:selenium binding activity; TAS.
DR   GO; GO:0006979; P:response to oxidative stress; TAS.
DR   Pfam; PF04593; SelP_C; 1.
DR   Pfam; PF04592; SelP_N; 1.
KW   Glycoprotein; Signal; Selenium; Selenocysteine; Plasma.
FT   SIGNAL        1     19
FT   CHAIN        20    381       SELENOPROTEIN P.
FT   DOMAIN      244    249       POLY-HIS.
FT   SE_CYS       59     59
FT   SE_CYS      300    300
FT   SE_CYS      318    318
FT   SE_CYS      330    330
FT   SE_CYS      345    345
FT   SE_CYS      352    352
FT   SE_CYS      367    367
```

```
FT   SE_CYS       369    369
FT   SE_CYS       376    376
FT   SE_CYS       378    378
FT   CARBOHYD      46     46       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      83     83       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     119    119       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     128    128       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     174    174       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     338    338       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   381 AA;  42705 MW;  1A82FFA0BCD13CD7 CRC64;

  Query Match           64.7%;  Score 33;  DB 1;  Length 381;
  Best Local Similarity  55.6%;  Pred. No. 50;
  Matches    5;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 KCLIQLFCK 9
              :|: || :|
Db        280 RCINQLLCK 288


RESULT 13
GTR2_RAT
ID   GTR2_RAT       STANDARD;      PRT;   522 AA.
AC   P12336;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Solute carrier family 2, facilitated glucose transporter, member 2
DE   (Glucose transporter type 2, liver).
GN   SLC2A2 OR GLUT2 OR GLUT-2.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89003066; PubMed=3048704;
RA   Thorens B., Sarkar H.K., Kaback H.R., Lodish H.F.;
RT   "Cloning and functional expression in bacteria of a novel glucose
RT   transporter present in liver, intestine, kidney, and beta-pancreatic
RT   islet cells.";
RL   Cell 55:281-290(1988).
CC   -!- FUNCTION: FACILITATIVE GLUCOSE TRANSPORTER. THIS ISOFORM LIKELY
CC       MEDIATES THE BIDIRECTIONAL TRANSFER OF GLUCOSE ACROSS THE PLASMA
CC       MEMBRANE OF HEPATOCYTES AND IS RESPONSIBLE FOR UPTAKE OF GLUCOSE
CC       BY THE BETA CELLS; MAY COMPRISE PART OF THE GLUCOSE-SENSING
CC       MECHANISM OF THE BETA CELL. MAY ALSO PARTICIPATE WITH THE
CC       NA(+)/GLUCOSE COTRANSPORTER IN THE TRANSCELLULAR TRANSPORT OF
CC       GLUCOSE IN THE SMALL INTESTINE AND KIDNEY.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: PRESENT IN LIVER, INTESTINE, KIDNEY AND
CC       BETA-PANCREATIC ISLET CELLS.
CC   -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. GLUCOSE
CC       TRANSPORTER SUBFAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; J03145; AAA41298.1; -.
DR   PIR; A31556; A31556.
DR   InterPro; IPR007114; MFS.
DR   InterPro; IPR005828; Sub_transporter.
DR   InterPro; IPR005829; Sug_transporter.
DR   InterPro; IPR003663; Sugar_transpt.
DR   Pfam; PF00083; sugar_tr; 1.
DR   PRINTS; PR00171; SUGRTRNSPORT.
DR   TIGRFAMs; TIGR00879; SP; 1.
DR   PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR   PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW   Transmembrane; Sugar transport; Transport; Glycoprotein;
KW   Multigene family.
FT   DOMAIN        1     10       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     11     31       1 (POTENTIAL).
FT   DOMAIN       32     96       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     97    117       2 (POTENTIAL).
FT   DOMAIN      118    125       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    126    146       3 (POTENTIAL).
FT   DOMAIN      147    156       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    157    177       4 (POTENTIAL).
FT   DOMAIN      178    185       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    186    206       5 (POTENTIAL).
FT   DOMAIN      207    215       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    216    236       6 (POTENTIAL).
FT   DOMAIN      237    301       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    302    322       7 (POTENTIAL).
FT   DOMAIN      323    337       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    338    358       8 (POTENTIAL).
FT   DOMAIN      359    365       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    366    386       9 (POTENTIAL).
FT   DOMAIN      387    401       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    402    422       10 (POTENTIAL).
FT   DOMAIN      423    431       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    432    452       11 (POTENTIAL).
FT   DOMAIN      453    459       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    460    480       12 (POTENTIAL).
FT   DOMAIN      481    522       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD     62     62       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   522 AA;  57085 MW;  075AB81E56CF33F7 CRC64;

  Query Match           64.7%;  Score 33;  DB 1;  Length 522;
  Best Local Similarity  62.5%;  Pred. No. 68;
  Matches    5;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KCLIQLFC 8
              :||: |||
Db        230 QCLLLLFC 237


RESULT 14
GTR2_MOUSE
ID   GTR2_MOUSE     STANDARD;      PRT;   523 AA.
AC   P14246; Q9DBA7;
DT   01-JAN-1990 (Rel. 13, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Solute carrier family 2, facilitated glucose transporter, member 2
DE   (Glucose transporter type 2, liver).
GN   SLC2A2 OR GLUT2 OR GLUT-2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c; TISSUE=Liver;
RX   MEDLINE=90098776; PubMed=2602116;
RA   Suzue K., Lodish H.F., Thorens B.;
RT   "Sequence of the mouse liver glucose transporter.";
RL   Nucleic Acids Res. 17:10099-10099(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6; TISSUE=Liver;
RX   MEDLINE=89366666; PubMed=2771649;
RA   Asano T., Shibasaki Y., Lin J.L., Akanuma Y., Takaku F., Oka Y.;
RT   "The nucleotide sequence of cDNA for a mouse liver-type glucose
RT   transporter protein.";
RL   Nucleic Acids Res. 17:6386-6386(1989).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Liver;
RX   MEDLINE=21085660; PubMed=11217851;
```

```
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [4]
RP   SEQUENCE OF 384-496 FROM N.A.
RX   MEDLINE=92111400; PubMed=1765007;
RA   Hogan A., Heyner S., Charron M.J., Copeland N.G., Gilbert D.J.,
RA   Jenkins N.A., Thorens B., Schultz G.A.;
RT   "Glucose transporter gene expression in early mouse embryos.";
RL   Development 113:363-372(1991).
RN   [5]
RP   TISSUE SPECIFICITY.
RC   STRAIN=C57BL/6;
RX   MEDLINE=93170163; PubMed=1289053;
RA   Smith D.E., Gridley T.;
RT   "Differential screening of a PCR-generated mouse embryo cDNA library:
RT   glucose transporters are differentially expressed in early
RT   postimplantation mouse embryos.";
RL   Development 116:555-561(1992).
CC   -!- FUNCTION: FACILITATIVE GLUCOSE TRANSPORTER. THIS ISOFORM LIKELY
CC       MEDIATES THE BIDIRECTIONAL TRANSFER OF GLUCOSE ACROSS THE PLASMA
CC       MEMBRANE OF HEPATOCYTES AND IS RESPONSIBLE FOR UPTAKE OF GLUCOSE
CC       BY THE BETA CELLS; MAY COMPRISE PART OF THE GLUCOSE-SENSING
CC       MECHANISM OF THE BETA CELL. MAY ALSO PARTICIPATE WITH THE
CC       NA(+)/GLUCOSE COTRANSPORTER IN THE TRANSCELLULAR TRANSPORT OF
CC       GLUCOSE IN THE SMALL INTESTINE AND KIDNEY.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: IN EMBRYO, EXPRESSED IN ENDODERM LAYER OF
CC       YOLK SAC AND LIVER PRIMORDIUM.
CC   -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. GLUCOSE
CC       TRANSPORTER SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X16986; CAA34855.1; -.
DR   EMBL; X15684; CAA33719.1; -.
DR   EMBL; S77926; AAB20847.1; -.
DR   EMBL; AK005068; BAB23792.1; -.
DR   PIR; S06920; S06920.
DR   MGD; MGI:1095438; Slc2a2.
DR   InterPro; IPR007114; MFS.
DR   InterPro; IPR005828; Sub_transporter.
DR   InterPro; IPR005829; Sug_transporter.
DR   InterPro; IPR003663; Sugar_transpt.
DR   Pfam; PF00083; sugar_tr; 1.
DR   PRINTS; PR00171; SUGRTRNSPORT.
DR   TIGRFAMs; TIGR00879; SP; 1.
DR   PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR   PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW   Transmembrane; Sugar transport; Transport; Glycoprotein;
KW   Multigene family.
FT   DOMAIN        1     10       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     11     31       1 (POTENTIAL).
FT   DOMAIN       32     97       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     98    118       2 (POTENTIAL).
FT   DOMAIN      119    126       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    127    147       3 (POTENTIAL).
FT   DOMAIN      148    157       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    158    178       4 (POTENTIAL).
FT   DOMAIN      179    186       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    187    207       5 (POTENTIAL).
FT   DOMAIN      208    216       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    217    237       6 (POTENTIAL).
FT   DOMAIN      238    302       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    303    323       7 (POTENTIAL).
FT   DOMAIN      324    337       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    338    358       8 (POTENTIAL).
FT   DOMAIN      359    367       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    368    388       9 (POTENTIAL).
FT   DOMAIN      389    401       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    402    422       10 (POTENTIAL).
FT   DOMAIN      423    432       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    433    453       11 (POTENTIAL).
FT   DOMAIN      454    460       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    461    481       12 (POTENTIAL).
FT   DOMAIN      482    523       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD     62     62       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    106    106       G -> D (IN REF. 2).
FT   CONFLICT    181    181       T -> N (IN REF. 1).
FT   CONFLICT    327    327       S -> T (IN REF. 2).
FT   CONFLICT    351    351       F -> L (IN REF. 1).
FT   CONFLICT    432    432       P -> S (IN REF. 1).
SQ   SEQUENCE   523 AA;  57106 MW;  F12B5376CBDB5F63 CRC64;

  Query Match             64.7%;  Score 33;  DB 1;  Length 523;
  Best Local Similarity   62.5%;  Pred. No. 68;
  Matches    5;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KCLIQLFC 8
              :||: |||
Db        231 QCLLLLFC 238


RESULT 15
S21F_MACFA
ID   S21F_MACFA     STANDARD;      PRT;   602 AA.
AC   Q9GMU6;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Solute carrier family 21 member 14.
GN   SLC21A14.
OS   Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9541;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA   Suzuki Y., Sugano S., Hashimoto K.;
RT   "Isolation of full-length cDNA clones from macaque brain cDNA
RT   libraries.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Mediates the Na(+)-independent high affinity transport
CC       of organic anions such as the thyroid hormones thyroxine (T4) and
CC       rT3. Other potential substrates, such as triiodothyronine (T3),
CC       estradiol-17-beta-glucuronide, estrone-3-sulfate and
CC       sulfobromophthalein (BSP) are transported with much lower
CC       efficiency (By similarity).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO THE SLC21 FAMILY OF TRANSPORTERS.
```

```
CC   ---------------------------------------------------------------------- --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------- --
DR   EMBL; AB047629; BAB12153.1; -.
DR   InterPro; IPR004157; OATP_Cterm.
DR   InterPro; IPR004156; OATP_Nterm.
DR   Pfam; PF03137; OATP_C; 1.
DR   Pfam; PF03132; OATP_N; 1.
DR   TIGRFAMs; TIGR00805; oat; 1.
KW   Transmembrane; Transport; Ion transport; Glycoprotein.
FT   TRANSMEM      79     99       POTENTIAL.
FT   TRANSMEM     111    131       POTENTIAL.
FT   TRANSMEM     243    263       POTENTIAL.
FT   TRANSMEM     283    303       POTENTIAL.
FT   TRANSMEM     317    337       POTENTIAL.
FT   TRANSMEM     440    460       POTENTIAL.
FT   TRANSMEM     481    501       POTENTIAL.
FT   TRANSMEM     533    553       POTENTIAL.
FT   DOMAIN       142    205       SER-RICH.
FT   CARBOHYD     146    146       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     199    199       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     400    400       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     410    410       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     423    423       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     582    582       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   602 AA;  66738 MW;  1E02E84170F6C31D CRC64;

  Query Match             64.7%;  Score 33;  DB 1;  Length 602;
  Best Local Similarity   100.0%;  Pred. No. 78;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 IQLFCK 9
              ||||||
Db          9 IQLFCK 14


Search completed: October 14, 2003, 14:34:10
Job time : 8.54167 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 14:33:18 ; Search time 29.8125 Seconds
(without alignments)
77.903 Million cell updates/sec

Title: US-09-881-490-191
Perfect score: 51
Sequence: 1 KCLIQLFCK 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query % Match	Length	DB	ID	Description
1	40	78.4	157	10	Q9SPY9	Q9spy9 pisum sativ
2	40	78.4	895	5	Q8SSZ4	Q8ssz4 dictyosteli
3	39	76.5	319	13	Q9YH55	Q9yh55 gallus gall
4	39	76.5	798	11	Q8BUJ2	Q8buj2 mus musculu
5	39	76.5	875	11	Q8BY36	Q8by36 mus musculu
6	39	76.5	944	3	Q96WT2	Q96wt2 phoma betae
7	38	74.5	313	13	Q9PSJ5	Q9psj5 ictalurus p
8	38	74.5	554	16	Q988N4	Q988n4 rhizobium l
9	38	74.5	661	11	Q61889	Q61889 mus musculu
10	38	74.5	662	11	Q9WV57	Q9wv57 rattus norv
11	38	74.5	700	10	Q9LRV2	Q9lrv2 arabidopsis
12	38	74.5	752	11	Q9JI82	Q9ji82 mus musculu
13	38	74.5	2754	5	Q8I451	Q8i451 plasmodium
14	37	72.5	163	10	Q9SY46	Q9sy46 arabidopsis
15	37	72.5	217	13	P70015	P70015 xenopus lae
16	37	72.5	217	13	P70017	P70017 xenopus lae
17	37	72.5	217	13	P70027	P70027 xenopus lae
18	37	72.5	548	5	Q8IRA7	Q8ira7 drosophila
19	37	72.5	1199	5	Q8I397	Q8i397 plasmodium
20	36	70.6	192	11	Q9D2J1	Q9d2j1 mus musculu
21	36	70.6	243	4	Q9H4Y5	Q9h4y5 homo sapien
22	36	70.6	248	11	Q9D2S1	Q9d2s1 mus musculu
23	36	70.6	248	11	Q8K2Q2	Q8k2q2 mus musculu
24	36	70.6	248	11	Q8BW12	Q8bw12 mus musculu
25	36	70.6	311	11	Q8VFH3	Q8vfh3 mus musculu
26	36	70.6	476	5	Q8T3V4	Q8t3v4 drosophila
27	36	70.6	476	5	Q9V309	Q9v309 drosophila
28	36	70.6	516	10	Q8L6I5	Q8l6i5 oryza sativ
29	36	70.6	730	10	Q9FMF6	Q9fmf6 arabidopsis
30	36	70.6	2134	12	Q9YLS4	Q9yls4 avian encep
31	35	68.6	55	6	Q9GMM7	Q9gmm7 macaca fasc
32	35	68.6	196	12	O57013	O57013 grapevine l
33	35	68.6	211	4	Q96PJ7	Q96pj7 homo sapien
34	35	68.6	217	13	P70016	P70016 xenopus lae
35	35	68.6	236	6	Q9XS71	Q9xs71 bos taurus
36	35	68.6	251	4	Q96AR4	Q96ar4 homo sapien
37	35	68.6	259	11	Q8BHJ0	Q8bhj0 mus musculu
38	35	68.6	319	2	Q9RMD1	Q9rmd1 acinetobact
39	35	68.6	341	5	O18100	O18100 caenorhabdi
40	35	68.6	432	16	Q8ZR79	Q8zr79 salmonella
41	35	68.6	432	16	Q8Z8Q7	Q8z8q7 salmonella
42	35	68.6	462	10	Q9CA73	Q9ca73 arabidopsis
43	35	68.6	584	16	Q9ZCD7	Q9zcd7 rickettsia
44	35	68.6	604	10	Q8RZK3	Q8rzk3 oryza sativ
45	35	68.6	614	10	Q9FJY0	Q9fjy0 arabidopsis

ALIGNMENTS

RESULT 1
Q9SPY9
ID Q9SPY9 PRELIMINARY; PRT; 157 AA.
AC Q9SPY9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Putative NBS-LRR type disease resistance protein (Fragment).
GN RGA-G3A.
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Vicieae; Pisum.
OX NCBI_TaxID=3888;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. B880-221;
RA Timmerman-Vaughan G., Frew T., Weeden N.F.;
RT "Characterization and linkage mapping of R-gene analogous DNA
RT sequences in pea (Pisum sativum L.).";
RL Submitted (JAN-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF123703; AAD52719.1; -.
DR InterPro; IPR002114; HPr_SerP_site.
DR InterPro; IPR002182; NB-ARC.
DR Pfam; PF00931; NB-ARC; 1.
DR PROSITE; PS00589; PTS_HPR_SER; 1.
FT NON_TER 1 1
FT NON_TER 157 157
SQ SEQUENCE 157 AA; 17949 MW; 9FA495ECC7ABD946 CRC64;

Query Match 78.4%; Score 40; DB 10; Length 157;
Best Local Similarity 66.7%; Pred. No. 2.2;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KCLIQLFCK 9
|| :|||| :
Db 125 KCALQLFCQ 133

```
RESULT 2
Q8SSZ4
ID   Q8SSZ4      PRELIMINARY;      PRT;   895 AA.
AC   Q8SSZ4;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AX4;
RA   Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
RA   Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA   Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT   "Sequence and Analysis of Chromosome 2 of Dictyostelium.";
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC116330; AAL96724.1; -.
DR   InterPro; IPR006768; CwfJ_C_1.
DR   InterPro; IPR006767; CwfJ_C_2.
DR   Pfam; PF04677; CwfJ_N_1; 1.
DR   Pfam; PF04676; CwfJ_N_2; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   895 AA;  104485 MW;  94895D6A284E3384 CRC64;

  Query Match             78.4%;  Score 40;  DB 5;  Length 895;
  Best Local Similarity   77.8%;  Pred. No. 8.7;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KCLIQLFCK 9
              |||||:| |
Db        739 KCLIQMFAK 747


RESULT 3
Q9YH55
ID   Q9YH55      PRELIMINARY;      PRT;   319 AA.
AC   Q9YH55;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Olfactory receptor-like protein COR3'beta.
GN   COR3'BETA.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94140361; PubMed=8307571;
RA   Reitman M., Grasso J.A., Blumenthal R., Lewit P.;
RT   "Primary sequence, evolution, and repetitive elements of the Gallus
RT   gallus (chicken) beta-globin cluster.";
RL   Genomics 18:616-626(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Saitoh N., Pikaart M.J., Recillas-Targa F., Felsenfeld G.;
RT   "The 3' chromatin boundary of the chicken beta globin locus.";
RL   Submitted (NOV-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; L17432; AAD03349.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   319 AA;  35651 MW;  7381B123EFF250D6 CRC64;

  Query Match             76.5%;  Score 39;  DB 13;  Length 319;
  Best Local Similarity   85.7%;  Pred. No. 6.1;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 CLIQLFC 8
              ||||:||
Db         97 CLIQMFC 103


RESULT 4
Q8BUJ2
ID   Q8BUJ2      PRELIMINARY;      PRT;   798 AA.
AC   Q8BUJ2;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   VPS39/VAM6-like protein homolog (Fragment).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Heart;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK084743; BAC39271.1; -.
FT   NON_TER     798     798
SQ   SEQUENCE   798 AA;  91720 MW;  6D11A25DF46B35BB CRC64;

  Query Match             76.5%;  Score 39;  DB 11;  Length 798;
  Best Local Similarity   75.0%;  Pred. No. 13;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 CLIQLFCK 9
              |||||:|:
Db        610 CLIQLYCE 617


RESULT 5
Q8BY36
ID   Q8BY36      PRELIMINARY;      PRT;   875 AA.
AC   Q8BY36;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   VPS39/VAM6-like protein homolog.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Thymus;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK042253; BAC31205.1; -.
SQ   SEQUENCE   875 AA;  100724 MW;  9BA3B0061BBF38F6 CRC64;

  Query Match             76.5%;  Score 39;  DB 11;  Length 875;
  Best Local Similarity   75.0%;  Pred. No. 14;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 CLIQLFCK 9
              |||||:|:
Db        610 CLIQLYCE 617
```

```
RESULT 6
Q96WT2
ID   Q96WT2      PRELIMINARY;      PRT;   944 AA.
AC   Q96WT2;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Aphidicolan-16beta-ol synthase.
GN   ACS.
OS   Phoma betae.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina;
OC   Dothideomycetes et Chaetothyriomycetes incertae sedis; Phoma.
OX   NCBI_TaxID=137527;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Oikawa H., Toyomasu T., Toshima H., Ohashi S., Kawaide H., Kamiya Y.,
RA   Otsuka M., Shinoda S., Mitsuhashi W., Sassa T.;
RT   "Cloning and functional expression of cDNA encoding aphidicolan-
RT   16beta-ol syntase: A key enzyme responsible for formation of an
RT   unusual diterpene skelton in biosynthesis of aphidicolin.";
RL   J. Am. Chem. Soc. 2001:5154-5155(2001).
DR   EMBL; AB049075; BAB62102.1; -.
DR   InterPro; IPR001906; Terp_synth-like.
DR   Pfam; PF01397; Terpene_synth; 1.
SQ   SEQUENCE   944 AA;  106580 MW;  41C781606F236CE5 CRC64;

  Query Match             76.5%;  Score 39;  DB 3;  Length 944;
  Best Local Similarity   75.0%;  Pred. No. 14;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 CLIQLFCK 9
              ||| :|||
Db        911 CLIDMFCK 918


RESULT 7
Q9PSJ5
ID   Q9PSJ5      PRELIMINARY;      PRT;   313 AA.
AC   Q9PSJ5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Odorant receptor 8 (Olfactory receptor 8).
OS   Ictalurus punctatus (Channel catfish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;
OC   Ictaluridae; Ictalurus.
OX   NCBI_TaxID=7998;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93201590; PubMed=7916654;
RA   Ngai J., Dowling M.M., Buck L., Axel R., Chess A.;
RT   "The family of genes encoding odorant receptors in the channel
RT   catfish.";
RL   Cell 72:657-666(1993).
DR   EMBL; L09222; AAK14716.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   313 AA;  35552 MW;  0DA8FA0F6F700963 CRC64;

  Query Match             74.5%;  Score 38;  DB 13;  Length 313;
  Best Local Similarity   71.4%;  Pred. No. 9.5;
  Matches    5;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 CLIQLFC 8
              ||:|:||
Db         98 CLVQMFC 104


RESULT 8
Q988N4
ID   Q988N4      PRELIMINARY;      PRT;   554 AA.
AC   Q988N4;
DT   01-OCT-2001 (TrEMBLrel. 18, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein mll6665.
GN   MLL6665.
OS   Rhizobium loti (Mesorhizobium loti).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Phyllobacteriaceae; Mesorhizobium.
OX   NCBI_TaxID=381;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MAFF303099;
RX   MEDLINE=21082930; PubMed=11214968;
RA   Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA   Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA   Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA   Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA   Takeuchi C., Yamada M., Tabata S.;
RT   "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT   Mesorhizobium loti.";
RL   DNA Res. 7:331-338(2000).
DR   EMBL; AP003009; BAB52913.1; -.
DR   InterPro; IPR001466; Beta_lactamase.
DR   Pfam; PF00144; beta-lactamase; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   554 AA;  59616 MW;  8453E0738C56F67F CRC64;

  Query Match             74.5%;  Score 38;  DB 16;  Length 554;
  Best Local Similarity   75.0%;  Pred. No. 15;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KCLIQLFC 8
              ||||: ||
Db        495 KCLIEYFC 502


RESULT 9
Q61889
ID   Q61889      PRELIMINARY;      PRT;   661 AA.
AC   Q61889;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   MPS1 protein precursor (Fragment).
GN   MPEG1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ICR;
RX   MEDLINE=95195232; PubMed=7888681;
RA   Spilsbury K., O'Mara M.A., Wu W.M., Rowe P.B., Symonds G.,
RA   Takayama Y.;
RT   "Isolation of a novel macrophage-specific gene by differential cDNA
RT   analysis.";
RL   Blood 85:1620-1629(1995).
DR   EMBL; L20315; AAA73957.1; -.
DR   MGD; MGI:1333743; Mpeg1.
DR   InterPro; IPR001862; MAC_perforin.
DR   Pfam; PF01823; MACPF; 1.
DR   SMART; SM00457; MACPF; 1.
KW   Signal.
FT   NON_TER       1      1
FT   SIGNAL       <1     26       POTENTIAL.
SQ   SEQUENCE   661 AA;  73166 MW;  95D2B55E0DB0C814 CRC64;

  Query Match             74.5%;  Score 38;  DB 11;  Length 661;
```

```
  Best Local Similarity   55.6%;  Pred. No. 17;
  Matches    5;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KCLIQLFCK 9
              || :::|||
Db        434 KCTLKIFCK 442


RESULT 10
Q9WV57
ID   Q9WV57      PRELIMINARY;      PRT;   662 AA.
AC   Q9WV57;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   Mpg-1 protein.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley;
RA   Chen S., Garcia G.E., Xia Y., Wilson C.B., Ku G., Feng L.;
RT   "Immunoneutralization of the Macrophage Gene-1 Product Attenuates
RT   Renal Injury in Experimental Glomerulonephritis in WKY rats.";
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF156540; AAD38417.1; -.
DR   InterPro; IPR001862; MAC_perforin.
DR   Pfam; PF01823; MACPF; 1.
DR   SMART; SM00457; MACPF; 1.
SQ   SEQUENCE   662 AA;  73327 MW;  98BC5F14B27BEC61 CRC64;

  Query Match             74.5%;  Score 38;  DB 11;  Length 662;
  Best Local Similarity   55.6%;  Pred. No. 17;
  Matches    5;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KCLIQLFCK 9
              || :::|||
Db        435 KCTLKIFCK 443


RESULT 11
Q9LRV2
ID   Q9LRV2      PRELIMINARY;      PRT;   700 AA.
AC   Q9LRV2;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Selenium-binding protein-like.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Columbia;
RA   Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Columbia;
RX   MEDLINE=20277480; PubMed=10819329;
RA   Nakamura Y.;
RT   "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
RT   features of the regions of 4,504,864 bp covered by sixty P1 and TAC
RT   clones.";
RL   DNA Res. 7:131-135(2000).
DR   EMBL; AB028611; BAB01836.1; -.
DR   InterPro; IPR002885; PPR.
DR   Pfam; PF01535; PPR; 12.
DR   TIGRFAMs; TIGR00756; PPR; 8.

SQ   SEQUENCE   700 AA;  79194 MW;  1525A5DE7F242682 CRC64;

  Query Match             74.5%;  Score 38;  DB 10;  Length 700;
  Best Local Similarity   62.5%;  Pred. No. 18;
  Matches    5;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 CLIQLFCK 9
              |:|:|:||
Db        638 CMIELYCK 645


RESULT 12
Q9JI82
ID   Q9JI82      PRELIMINARY;      PRT;   752 AA.
AC   Q9JI82;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   EPCS50.
GN   PFPL OR EPCS50.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Hemberger M.C., Himmelbauer H., Ruschmann J., Zeitz C., Fundele R.H.;
RT   "cDNA subtraction cloning reveals novel genes whose temporal and
RT   spatial expression indicates association with trophoblast invasion.";
RL   Dev. Biol. 0:0-0(2000).
DR   EMBL; AF250839; AAF81276.1; -.
DR   MGD; MGI:1860266; Pfpl.
DR   InterPro; IPR001862; MAC_perforin.
DR   SMART; SM00457; MACPF; 1.
SQ   SEQUENCE   752 AA;  83921 MW;  C2B8ED4D871D894F CRC64;

  Query Match             74.5%;  Score 38;  DB 11;  Length 752;
  Best Local Similarity   44.4%;  Pred. No. 19;
  Matches    4;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KCLIQLFCK 9
              ||:::::||:
Db        438 KCILKIFCR 446


RESULT 13
Q8I451
ID   Q8I451      PRELIMINARY;      PRT;  2754 AA.
AC   Q8I451;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein.
GN   PFE0250W.
OS   Plasmodium falciparum (isolate 3D7).
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=36329;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Devlin K., Baker S., Davies P., Mungal K., Berriman M., Pain A.,
RA   Hall N., Bowman S., Churcher C., Quail M., Barrell B.;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22255708; PubMed=12368867;
RA   Hall N., Pain A., Berriman M., Churcher C., Harris B., Harris D.,
RA   Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,
RA   Buckee C.O., Burrows C., Cherevach I., Chillingworth C.,
RA   Chillingworth T., Christodoulou Z., Clark L., Clark R., Corton C.,
RA   Cronin A., Davies R., Davis P., Dear P., Dearden F., Doggett J.,
RA   Feltwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N., Hance Z.,
RA   Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P.,
RA   Humphray S., Jagels K., James K.D., Johnson D., Kerhornou A.,
```

```
RA   Knights A., Konfortov B., Kyes S., Larke N., Lawson D., Lennard N.,
RA   Line A., Maddison M., Mclean J., Mooney P., Moule S., Murphy L.,
RA   Oliver K., Ormond D., Price C., Quail M.A., Rabbinowitsch E.,
RA   Rajandream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M.,
RA   Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K.,
RA   Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,
RA   Sulston J.E., Craig A., Newbold C., Barrell B.G;
RT   "Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13.";
RL   Nature 419:527-531(2002).
DR   EMBL; AL929351; CAD51416.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   2754 AA;  328673 MW;  44FB731EF95DE073 CRC64;

  Query Match             74.5%;  Score 38;  DB 5;  Length 2754;
  Best Local Similarity   55.6%;  Pred. No. 53;
  Matches    5;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KCLIQLFCK 9
              |||:: :|||
Db       2022 KCMMTMFCK 2030


RESULT 14
Q9SY46
ID   Q9SY46      PRELIMINARY;      PRT;   163 AA.
AC   Q9SY46;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Hypothetical 18.6 kDa protein.
GN   T5L23.18 OR AT4G03680.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RA   Zhong J., Ma P., Parnell L.D., Chen C.-N., Chen E.Y.;
RT   "BAC T5L23 from chromosome IV, position 19 cM.";
RL   Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Zhong J., Ma P., Parnell L.D., Chen C.N., Chen E.Y., Mewes H.W.,
RA   Lemcke K., Mayer K.F.X.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC005142; AAD15324.1; -.
DR   EMBL; AL161497; CAB77853.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   163 AA;  18568 MW;  658F2CE68C39EF4F CRC64;

  Query Match             72.5%;  Score 37;  DB 10;  Length 163;
  Best Local Similarity   100.0%;  Pred. No. 9;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KCLIQLF 7
              |||||||
Db          7 KCLIQLF 13


RESULT 15
P70015
ID   P70015      PRELIMINARY;      PRT;   217 AA.
AC   P70015;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Olfactory receptor (Fragment).
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Freitag J.;
RL   Submitted (SEP-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96112032; PubMed=8845161;
RA   Freitag J., Krieger J., Strotman J., Breer H.;
RT   "Two classes of olfactory receptors in Xenopus laevis.";
RL   Neuron 15:1383-1392(1995).
DR   EMBL; Y08346; CAA69632.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
FT   NON_TER       1      1
FT   NON_TER     217    217
SQ   SEQUENCE   217 AA;  24780 MW;  2590F8AE177CDFE2 CRC64;

  Query Match             72.5%;  Score 37;  DB 13;  Length 217;
  Best Local Similarity   85.7%;  Pred. No. 11;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 CLIQLFC 8
              || |||:
Db         27 CLFQLFC 33


Search completed: October 14, 2003, 14:37:00
Job time : 31.8125 secs
```

PALM Pay Periods for FY03

PP	Beginning Date	Ending Date	
0301	October 01, 2002	October 19, 2002	
0302	October 20, 2002	November 02, 2002	
0303	November 03, 2002	November 16, 2002	
0304	November 17, 2002	November 30, 2002	
0305	December 01, 2002	December 14, 2002	End of 1st Quarter
0306	December 15, 2002	December 28, 2002	
0307	December 29, 2002	January 11, 2003	
0308	January 12, 2003	January 25, 2003	
0309	January 26, 2003	February 08, 2003	
0310	February 9, 2003	February 22, 2003	
0311	February 23 2003	March 08, 2003	
0312	March 9, 2003	March 22, 2003	End of 2nd Quarter
0313	March 23, 2003	April 05, 2003	
0314	April 06, 2003	April 19, 2003	
0315	April 20, 2003	May 03, 2003	
0316	May 04, 2003	May 17, 2003	
0317	May 18, 2003	May 31, 2003	
0318	June 01, 2003	June 14, 2003	
0319	June 15, 2003	June 28, 2003	End of 3rd Quarter
0320	June 29, 2003	July 12, 2003	
0321	July 13, 2003	July 26, 2003	
0322	July 27, 2003	August 9, 2003	
0323	August 10, 2003	August 23, 2003	
0324	August 24, 2003	September 06, 2003	
0325	September 07, 2003	September 20, 2003	
0326	September 21, 2003	September 30, 2003	End of 4th Quarter